Dysregulation of histone acetylation and decoupling from transcription in Parkinson's disease

Lilah Toker, PhD[1,2†], Gia T Tran, MD[1,2†], Janani Sundaresan, MSc[1,2], Ole-Bjørn Tysnes, MD, PhD[1,2], Guido Alves, MD, PhD[3,4], Kristoffer Haugarvoll, MD, PhD[1,2], Gonzalo S Nido, PhD[1,2], Christian Dölle, PhD[1, 2] and Charalampos Tzoulis, MD, PhD[1,2*]

1. Neuro-SysMed, Department of Neurology, Haukeland University Hospital, 5021 Bergen, Norway. 2. Department of Clinical Medicine, University of Bergen, Pb 7804, 5020 Bergen, Norway. 3. The Norwegian Centre for Movement Disorders and Department of Neurology, Stavanger University Hospital, Pb 8100, 4068 Stavanger, Norway. 4. Department of Mathematics and Natural Sciences, University of Stavanger, 4062 Stavanger, Norway.

[†] These authors contributed equally to the work

[*]**Correspondence to:**
Prof. Charalampos Tzoulis
Neuro-SysMed Center of Excellence for Clinical Research in Neurological Diseases
Department of Neurology, Haukeland University Hospital
Department of Clinical Medicine, University of Bergen
5021 Bergen, Norway.
Telephone: +47 55975061
E-mail: charalampos.tzoulis@nevro.uib.no
E-mail-2: charalampos.tzoulis@helse-bergen.no

Number of words in body of manuscript: 4,619

Number of display items: 7

Keywords: H3K27, epigenetics, neurodegeneration, ChIP-seq, sirtuin, mitochondria, gene expression

List of abbreviations:

ChIP-seq: chromatin immunoprecipitation sequencing

Ctrl: control

DAR: differentially acetylated region

EGL: effective gene length

GWAS: genome-wide association study

HAT: histone acetyltransferase

HDAC: histone deacetylase

L-DOPA: levodopa

MSPs: Marker Site Profiles

NAD: nicotinamide adenine dinucleotide

NBB: Netherlands Brain Bank

PD: Parkinson's disease

PW: Park West

RiP: reads in peaks

RLE: Relative Log Expression

SIRT: sirtuin

SNc: substantia nigra pars compacta

WB: Western blot

WES: whole exome sequencing

Abstract

The etiology of Parkinson's disease (PD) is mostly unaccounted for. Based on previous evidence of neuronal metabolic defects in PD, we hypothesized that altered histone acetylation may be implicated in the disease process. Here, we present the first genome-wide study of histone acetylation in PD. Our findings show hyperacetylation of multiple histone sites in the PD prefrontal cortex, with the most prominent change observed for the H3K27 site, a marker of active promoters and enhancers. Chromatin immunoprecipitation sequencing analysis indicates that H3K27 hyperacetylation occurs throughout the genome but is overrepresented among genes previously associated with PD. Intriguingly, we find strong evidence that H3K27 acetylation is decoupled from gene expression in PD. Our results replicate across two independent cohorts and establish aberrant histone acetylation as an important element of the molecular pathophysiology of PD.

Parkinson's disease (PD) is the second most common neurodegenerative disorder, affecting ~1.8 % of the population above the age of 65 years[1]. The neuropathological hallmark of PD is the progressive loss of the dopaminergic neurons of the the *substantia nigra pars compacta* (SNc) in the presence of α-synuclein-positive inclusions termed Lewy pathology. Additional neurodegenerative changes occur in multiple regions of the nervous system, including several brainstem nuclei, the olfactory bulb, hippocampus, amygdala and the neocortex, as well as the autonomic and enteric nervous systems[2]. At the cellular level, PD is characterized by altered mitochondrial function, including complex-I deficiency, which has been shown to occur throughout the brain[3,4].

While monogenic forms exist[5], the vast majority of patients have "idiopathic PD", which is of complex etiology. Genome-wide association studies (GWAS) have identified multiple risk loci associated with idiopathic PD[6], but these make a relatively small collective contribution to the disease risk. Moreover, these have not been linked to specific molecular mechanisms that could be targeted therapeutically.

Epigenetic modifications, such as DNA methylation and histone acetylation, have an emerging role in modulating the risk and progression of complex disorders. Dysregulation of histone acetylation has been previously linked to aging[10] and neurodegeneration[11] and changes in histone acetylation at specific genomic regions have been reported in Alzheimer's disease[12,13]. In PD, only a few small-scale studies of histone acetylation have been reported, providing inconclusive results[14–16]. Moreover, the genomic landscape of histone acetylation in PD has never been assessed.

While the question of aberrant histone acetylation in PD remains largely unexplored, available evidence suggests that it may be involved in the pathophysiology of the disease. Histone acetylation is tightly coupled to mitochondrial function via the activity of the NAD^+-dependent class III histone deacetylases known as sirtuins[17]. Mitochondrial complex-I

deficiency decreases the NAD^+/NADH ratio, which can lead to histone hyperacetylation via decreased sirtuin activity[18–21]. Indeed, decreased sirtuin activity and hyperacetylation of several histone sites including H3K9, H3K27 and H2BK15[15] have been observed following treatment of neuronal cells with PD-associated mitochondrial complex-I inhibitors[15,18,22]. Moreover, reduction in NAD^+ availability, decreased sirtuin function and histone hyperacetylation have all been shown to occur during aging[10,21], which is a prerequisite for idiopathic PD.

Here, we report the first genome-wide study of histone acetylation in the PD brain, based on two independent cohorts from the Norwegian ParkWest study (PW)[23] and the Netherlands Brain Bank (NBB). Our findings indicate hyperacetylation of multiple histone sites and reveal that H3K27 acetylation is profoundly dysregulated at a genome-wide level and decoupled from transcription in the PD brain.

Results

Increased histone acetylation and histone protein levels in the PD brain

Our first goal was to determine whether the global acetylation landscape is altered in the PD brain. To this end, we performed an exploratory Western blot analysis of protein acetylation in the prefrontal cortex of 13 individuals with PD and 13 demographically matched controls, using an antibody against acetylated lysine (Fig. 1a, Supplementary Table S1). We concentrated on the prefrontal cortex because, while it is affected by the disease, it is less likely to be biased by changes in cell composition, such as neuronal loss and glial infiltration, which confound observations in severely affected areas such as the *substantia nigra*[2]. Linear regression analysis of the densitometric measurements adjusted for age, sex and postmortem interval (PMI) indicated an increased acetylation signal at 17 kDa, consistent with the molecular weight of core histone proteins ($p = 0.014$, Fig. 1a).

To confirm that the hyperacetylated proteins were indeed histones and to dissect histone- and site-specificity, we performed an in-depth analysis of specific acetylation sites across the four core histones: H2AK5ac, H2BK15ac, H3K9/K14ac, H3K27ac, H3K56ac, H4K5ac, H4K12ac, and H4K16ac. The acetylation status of α-tubulin, a cytoskeletal protein typically regulated by acetylation, was assessed as a non-histone control. To maximize sensitivity, we focused our analyses on the 7 PD and 7 control samples showing the strongest difference in the pan-acetylation signal (Supplementary Table S1). The acetylated fraction of each histone site (the ratio between the signal densities of the acetylated histone and the total histone) as well as total histone quantity were compared between PD and controls using linear mixed models with gender, age and PMI as fixed effects and subject and blot number as random effects. These analyses indicated a significant increase in H2BK15ac ($p = 1.4 \times 10^{-5}$), H3K9/14ac ($p = 3.9 \times 10^{-4}$), H3K27ac ($p = 4.5 \times 10^{-8}$), H3K56ac ($p = 0.01$) and H4K12ac ($p = 5.0 \times 10^{-3}$) in PD. No significant difference was found in the acetylated fraction of H2AK5,

H4K5, H4K16, or α-tubulin (Fig. 1b,c, Supplementary Table S2). Interestingly, the total quantity of the four core histone proteins was also elevated in PD (H2A: $p = 2.7 \times 10^{-5}$, H2B: $p = 1.1 \times 10^{-8}$, H3: $p = 2.0 \times 10^{-7}$; H4: $p = 3.9 \times 10^{-18}$).

While we did not expect major changes in cellular composition in the prefrontal cortex of PD patients, we wanted to rule out this potential bias in our samples. To this end, we carried out immunoblot analyses of several cell-type markers: NeuN (neurons), GFAP (astrocytes), CNP1 (oligodendrocytes) and CX3CR1 (microglia) as well as the nuclear marker NUP188, in the same samples used for acetylation assessment. The only significant change was observed for CNP1, showing increased signal intensity in PD samples ($p = 3.9 \times 10^{-5}$, Fig. 1d, Supplementary Table S2). To ensure that our findings were not driven by increased oligodendrocyte abundance in the PD samples, we repeated our prior immunoblot analyses adding CNP1 as a covariate to the statistical model. As shown in Supplementary Table S2, all of the indicated hyperacetylated sites remained significant after CNP1 adjustment.

In order to determine the anatomical specificity of H3K27ac and H3K9/14ac in the PD brain, we assessed these two markers in the cerebellum and striatum from the same patients. H3K27 was significantly hyperacetylated in the PD striatum ($p = 0.034$) and cerebellum ($p = 0.016$). H3K9/14 was significantly hypoacetylated in the PD cerebellum ($p = 7.0 \times 10^{-3}$) and unchanged in the striatum ($p = 0.7$, Fig. 2, Supplementary Table S2).

Histone hyperacetylation in the PD prefrontal cortex is accompanied by upregulation of sirtuin protein levels

Our study was motivated by the hypothesis that the well documented complex-I deficiency in the PD brain is likely to decrease the NAD^+/NADH ratio, thereby reducing the activity of sirtuins, the NAD^+-dependent histone deacetylases. A direct confirmation of this hypothesis would require a reliable determination of intracellular NAD^+ levels in our samples, which is

not feasible due to its rapid degradation in postmortem tissue[24]. However, it was previously reported that decreased sirtuin activity caused by NAD^+ deficiency is accompanied by increase in the protein quantity[25,26]. Thus, we assessed whether sirtuin protein levels were altered in the samples tested for histone acetylation. We focused on SIRT1, SIRT2 and SIRT3, localized to the nucleus, cytosol and mitochondria, respectively, since they have been linked to aging and neurodegeneration[17]. In line with our hypothesis, immunoblot analysis, indicated a significant upregulation of SIRT1 ($p = 2.0 \times 10^{-3}$) and SIRT3 ($p = 3.0 \times 10^{-3}$) in PD samples. No difference was detected for SIRT2 ($p = 0.57$, Fig. 2c, Supplementary Table 2).

Histone hyperacetylation in PD samples is not induced by anti-Parkinson drugs in-vitro

We next wondered whether the observed altered histone acetylation in PD is likely to result from treatment with anti-Parkinson agents. To test this, we exposed fully differentiated SH-SY5Y cells to the drugs used by most individuals with PD from the PW cohort during their last year of life (L-DOPA, carbidopa, entacapone and ropinirole), and monitored acetylation of H3K9/14 and H3K27. Immunoblot analysis revealed no effect for any of the tested anti-Parkinson drugs on histone acetylation (Fig. 2d).

Acetylated H3K27 exhibits higher genomic occupancy in PD

Our immunoblot analysis pointed out H3K27ac as the modification showing the most pronounced alteration in the PD prefrontal cortex. Our next step was to gain insight into the genomic regions undergoing this alteration. To this end, we carried out ChIP-seq analysis of H3K27ac in samples from the PW cohort, henceforth referred to as the discovery cohort (controls = 11, PD = 17, Supplementary Table S1).

Our first goal was to characterize the genomic distribution of the H3K27ac mark in PD and control samples. For this purpose, we identified H3K27ac enriched genomic regions (peaks) in each group separately (see methods). In line with the increased H3K27ac signal observed in the immunoblot analysis, higher number of peaks (146,763 vs. 135,662), as well as a higher proportion of group unique peaks (32.2% vs. 13.2%), and wider genome coverage (10.6% vs. 9.11%) were identified in PD compared to controls (Fig. 3a, Supplementary Table 3). We then compared the distribution of genomic annotations and detection p-values of the common (overlapping) and the unique peaks between the PD and control groups. In both groups, unique peaks occurred more frequently in intergenic and intronic regions than in promoter and exonic regions (Supplementary Fig. S1a, b). One possible interpretation for this observation is that unique peaks show a predilection for distal enhancer regions. Notably, however, the unique peaks were also characterized by significantly higher detection p-values, compared to the common peaks (Fig. S1c, d). Thus, we believe that a more plausible explanation is that the majority of the unique peaks represent random localization of acetylated H3K27.

ChIP-seq data corroborates H3K27 hyperacetylation in the PD brain

Next, we redefined H3K27ac peaks by identifying H3K27ac enriched regions in aggregated samples (see methods). The final peak-set consisted of 160,521 peaks, covering 11.6% of the genome and distributed over all 22 autosomes and both sex chromosomes (median peak size, 817bp, Fig. 3d). All downstream analyses in the discovery cohort were performed based on this peak-set, further referred to as PW_peak-set.

ChIP-seq data reflects the abundance and the genomic distribution of the target protein (or protein modification) which can vary substantially between samples[27]. Thus, the inter-individual variation in ChIP-seq library size represents a combination of noise and genuine

biological variation. In our study, we anticipated an increase in overall H3K27ac ChIP-seq signal in PD, based on the observed increase in H3K27ac fraction in the Western blot analysis. We therefore reasoned that the most suitable approach for normalization of our data would be using reads in peaks (RiP) annotated to house-keeping genes[27]. We modified this approach by manually selecting peaks annotated to known housekeeping genes and exhibiting a high and broad peak signal covering most of the gene body. The rational here was that these genes should be saturated with histone acetylation and are, therefore, expected to exhibit little biological variation. Namely, any observed variation in these peaks is likely to represent noise. Supplementary Figure S2 shows the H3K27ac ChIP-seq peaks annotated to the previously proposed[27] and the chosen housekeeping genes.

To validate our normalization approach, we performed additional H3K27ac immunoblot and ChIP-seq analyses in samples from 55 individuals (including the subjects from our discovery cohort, Supplementary Table S1) and tested the correlation between the immunoblot and ChIP-seq data normalized by different approaches. As shown in Supplementary Figure S3, the highest correlation between immunoblotting and ChIP-seq data was observed when the total number of RiP was normalized to our manually chosen peaks. Importantly, comparison of normalized RiP between PD and control samples corroborated the finding of general H3K27 hyperacetylation in PD (Supplementary Fig. S3).

Cell composition is a major source of variation in bulk tissue H3K27ac ChIP-seq data

It is established that different cell-types exhibit distinct epigenetic landscapes[28] and therefore, bulk tissue data must be evaluated in a context of its underlying cellular composition[29–31]. In order to estimate the relative cellular composition in our ChIP-seq data, we used an approach similar to the one described by Mancarci et al.[30]. Briefly, we intersected genomic H3K27ac binding sites differentiating between NeuN-positive and NeuN-negative brain cells[28] with the

list of brain cell-type specific marker-genes[30] to obtain genomic H3K27ac cell-type marker sites. The first principal component of reads aligned to the marker sites was used to obtain Marker Site Profiles (MSPs) which serve as a proxy for the relative abundance of the corresponding cell-types across the samples. We validated this approach in ChIP-seq data from cortical and cerebellar bulk tissue data[32], successfully recapitulating the well-known increased abundance of glial cells in cortical samples[33,34] (Supplementary Fig. S4).

Next, we performed MSP analysis in samples from our discovery cohort. In contrast to our findings from the immunoblot analysis, we found no significant group differences in any of the MSPs (Supplementary Figure S5). This is hardly surprising, since the immunoblot and ChIP-seq analyses were carried out in adjacent, but different tissue samples. Moreover, only a small subset of individuals included in ChIP-seq analysis was tested by immunoblotting. Since no difference was found between PD and controls, we reasoned that the cellular composition of our cortical samples would be mostly affected by white matter contamination that can be approximated in our data by oligodendrocyte MSPs.

We next assessed the effect of oligodendrocyte MSPs and other experimental and demographic covariates on our ChIP-seq data. Strikingly, sample-to-sample correlation of the PW_peak-set data showed that the samples clustered mainly based on oligodendrocytes MSPs (Fig. 3b). Accordingly, the first principal component of the data was mostly associated with oligodendrocyte MSP (Supplementary Figure S6), indicating that cellular composition is a major source of variance in our data, and therefore, oligodendrocyte MSPs were accounted for in all downstream analyses.

Hyperacetylation of H3K27 in the PD brain is a genome-wide phenomenon

Next, we sought to identify differentially acetylated regions (DARs) for H3K27 between PD and control samples in our discovery cohort. To further constrain our analysis to high

confidence regions, we only included peaks with an average coverage >5 counts/base (see methods). After filtering, 133,716 peaks annotated to 17,182 genes remained for downstream analyses. Five samples (one control and four PD samples) were detected as outliers and excluded from differential peak analysis (Supplementary Table S1). Differential peak analysis was performed using the "DESeq2" R package, adjusting for age, sex, batch, PMI, oligodendrocyte MSP and normalization peak ratio. DARs were defined as H3K27ac binding sites (peaks) with adjusted p-value < 0.05.

Our analysis identified 2,876 hyperacetylated and only 14 hypoacetylated DARs in PD, annotated to 1,432 and 9 genes, respectively. An MA plot (the log ratio vs. mean count) of all regions included in the analysis indicated a general trend for H3K27 hyperacetylation in PD, independent of the number of reads aligned to a region (Fig. 3c), strongly suggesting a genome-wide hyperacetylation of H3K27.

In light of the apparent general H3K27 hyperacetylation, the 14 differentially hypoacetylated regions are of particular interest since they are more likely to represent specifically targeted or regulated modifications. These regions were annotated to the following nine genes: *PTPRH*, *FAT1*, *ADAMTS9*, *ADAMTS9-AS*, *IL15RA*, *SLFN5*, *SLC11A1*, *JUP* and *PLEKHG1* (the same DAR was annotated to both *ADAMTS9*, *ADAMTS9-AS*). In five of these (*PTPRH*, *ADAMTS9*, *ADAMTS9-AS*, *IL15RA* and *SLFN5*) the DAR was localized to the promoter region of the gene. The full list of DARs with their associated genes is provided in Supplementary Table S3.

H3K27 hyperacetylation in the PD cortex is replicated in an independent cohort of subjects

In order to validate our findings from the discovery cohort, we carried out ChIP-seq analysis in an independent cohort of subjects (11 controls, 10 PD) from the Netherlands Brain Bank

(NBB, Supplementary Table S1). Compared to the discovery cohort, we identified overall less H3K27ac peaks and a slightly smaller genomic coverage in NBB (109,183 peaks after filtering, annotated to 17,306 genes, Supplementary Table S3). Nevertheless, when peaks were called for each group separately, we observed an increase in the percentage of unique peaks as well as increased total genomic coverage in the PD group (Supplementary table S3), confirming the trend for H3K27 hyperacetylation observed in our discovery cohort. The distribution of genomic annotations among the unique and common peaks and the distribution of peak p-values were similar between the two cohorts (Supplementary Fig. S1). Similar to the discovery cohort, cellular composition accounted for most of the interindividual variance in this cohort (Supplementary Fig. S6b, S7a).

Differential acetylation analysis confirmed our main findings from the discovery cohort, namely, genome-wide H3K27 hyperacetylation in the PD group (Supplementary Figure S5b). In total, the analysis indicated 2,486 hyperacetylated DARs, mapped to 937 genes, and 227 hypoacetylated DARs, mapped to 252 genes (Supplementary table S5). Strikingly, 272 out of the 937 genes with hyperacetylated DARs and 2 out of 252 genes with hypoacetylated DARs replicated across both cohorts (hypergeometric test: $p = 1.2 \times 10^{-82}$ and $p = 6.7 \times 10^{-3}$, respectively).

We next wondered whether genes previously associated with PD were over-represented among the genes replicating between the two cohorts. Among the 88 genes associated with PD based on the most recent GWAS metaanalysis[6], 75 could be assessed in our ChIP-seq data. Five out of the 274 replicating genes (272 hyperacetylated and two hypoacetylated): *DLG2*, *MAP4K4*, *CRHR1*, *MBNL2*, and *SH3GL2* were among the GWAS hits (hypergeometric p-value = 1.7×10^{-3}).

Next, we sought to determine to which extent the concordance between the two cohorts observed at the gene level is restricted to specific genomic locations. In order to harmonize

the genomic regions identified as peaks in the two cohorts, we focused on peaks with the same structural and functional genomic annotation (e.g. promoter, exon, 5'UTR) in both cohorts. These are henceforth referred to as common peaks. For example, two peaks were considered common if both were annotated to the first exon of the *PTPRH* gene, but not if one was annotated to the first exon of *PTPRH* while the other to the second exon of this gene. For each genomic region, we then calculated the Fisher's meta p-value for the pair of common peaks, termed metaP. Out of 48,008 pairs of common peaks, 2,756 and 2,893 were independently identified as DARs in the NBB and PW cohorts, respectively, with 288 DARs identified in both (hypergeometric p-value = 2.97×10^{-24}). These 288 DARs are henceforth referred to as replicating DARs. Table 1 summarizes the top replicating DARs (ranked based on metaP). The full list of replicating DARs and their genomic annotations is provided as Supplementary Table S6. Of note, 22 out of the 75 represented GWAS hits[6], had common peaks with adjusted metaP < 0.05 (hypergeometric p-value = 0.02, Supplementary Table S6). These findings suggest that differential H3K27 acetylation is overrepresented among genes previously associated with PD.

Data normalization approach and adjustment for cell-type composition greatly impact H3K27ac ChIP-seq analysis.

We next evaluated the impact of data normalization and cell-type correction on the output of our ChIP-seq analysis. For this purpose, we repeated the analysis with data normalized using Relative Log Expression (RLE) method, implemented in the DESeq2 package and performed with or without MSP adjustment. As expected, when data was normalized using the RLE method, which ensures that the median fold of change of all signals in a sample relative to all samples is zero, the calculated fold change was centered around zero (Fig. 4c,d). Interestingly, based on this normalization, most DARs exhibited downregulation in PD.

Oligodendrocyte MSP correction dramatically increased the number of DARs in both cohorts without having a big effect on the direction of change. A plausible explanation to this phenomenon is reduction in the sample-sample variability induced by cellular composition differences (Fig. 3b**,** Supplementary Fig. S6, S7a). Of notice, several genes robustly exhibited replicating DARs regardless of the normalization method (Supplementary Table S7). Figure 4 shows examples for two of such genes, *PTPRH* and *DLG2*, both of which have previously been associated with PD based on genetic studies[6,35–37] and were found in the top list of common DARs (Table 1).

Differential H3K27 acetylation in the PD brain is associated with the effective gene length and is more prevalent among PD-associated genes.

Our top replicating DARs included a large number of genes that are among the longest in the genome. We reasoned that this may be because longer genes have higher histone binding capacity and therefore are more likely to show a differential signal in the event of a generalized H3K27 hyperacetylation. To test our hypothesis, we assessed whether altered H3K27 acetylation, defined as -log10(metaP), is associated with the effective gene length (EGL), defined as the sum of all regions annotated to a gene (see methods). As shown in Figure 5a, we observed a positive association between EGL and the -log10(metaP) value for hyperacetylated regions, and a negative association for hypoacetylated regions.

Interestingly, we noted that PD-GWAS gene hits are also characterized by longer EGL. Thus, we wondered whether the observed over-representation of GWAS hits among the DARs was confounded by their EGL. To this end, we assessed the association between the region's -log10(metaP) and being annotated to a GWAS gene while adjusting for EGL using a linear model. The total number of common regions in a gene was included as a covariate in the model (see methods section). For hyperacetylated regions, the analysis indicated a

significant positive association between altered acetylation and being a GWAS-related gene as well as the EGL (GWAS: $\beta = 0.33$, $p = 4.3 \times 10^{-6}$; log10(EGL): $\beta = 0.42$, $p < 2 \times 10^{-16}$). For the hypoacetylated regions, the analysis indicated a significant negative association with EGL ($\beta = -0.06$, $p = 8.0 \times 10^{-4}$), and no association with being a GWAS-hit ($p = 0.3$). Analysis of each cohort independently corroborated the findings based on common regions (Figure 5b, Supplementary Table S8). These results indicate that while differential H3K27ac hyperacetylation exhibited a strong predilection for longer genes in the PD brain, this cannot fully explain the overrepresentation of PD-associated genes among the hyperacetylated DARs.

Promoter H3K27 acetylation is decoupled from gene expression in PD

It is well accepted that H3K27ac is a marker of active promoters and enhancers[38] and that modifications of this mark are generally predictive of gene expression[39]. Therefore, the global H3K27 hyperacetylation in the PD prefrontal cortex is predicted to cause a widespread increase in gene expression. Since such a phenomenon has not been reported, we wondered whether H3K27 acetylation and gene transcription may be decoupled in the PD brain.

To address this question, we analyzed RNA-seq data from prefrontal cortex of the same individuals included in the ChIP-seq analysis (Supplementary Table 1). A detailed description of the RNA-seq data analysis was recently published[40]. For each gene, we calculated Pearson's correlation between H3K27ac state (ChIP-seq) and transcript levels (RNA-seq) across control or PD individuals. We concentrated on peaks annotated to promoter regions of the genes, since these are expected to highly correlate with gene expression. For both ChIP-seq and RNA-seq data the correlation was performed between covariate-corrected signals. The correction covariates were based on the models used for differential analysis and included age, sex, PMI, batch and cellular composition.

As expected, the distribution of correlations between promoter H3K27ac and gene expression in control samples was shifted towards positive values (Fig. 6). However, in both cohorts the distribution of correlations in the PD group was centered significantly closer to 0. Remarkably, this remained true regardless of the correlation threshold chosen. While as expected, a higher threshold for absolute correlation resulted in increased median correlation in controls, this effect was much smaller for the PD group, resulting in bigger differences between the groups. Importantly, similar results were obtained based on RLE-normalized ChIP-seq data (Supplementary Fig. S8), indicating that the observed decreased correlation between H3K27ac and transcript levels in PD was not driven by our normalization method. These results indicate that, the coupling between promoter H3K27 acetylation state and the proximal gene transcription is impaired in in the PD prefrontal cortex.

Discussion

Our study addresses the pertinent question of whether aberrant histone acetylation occurs in the brain of individuals with PD. We report hyperacetylation of multiple histone sites in the PD prefrontal cortex with the most pronounced changes observed for H3K27, a marker of active promoters and enhancers. Our findings, replicated in two independent cohorts, indicate that H3K27 hyperacetylation is a genome-wide phenomenon, rather than being restricted to specific genomic regions. Intriguingly, our results also strongly suggest decoupling between H3K27 acetylation at promoter regions and gene expression in PD.

We were inspired by the accumulating evidence of mitochondrial dysfunction, suggesting that a hyperacetylation-prone cellular environment may exist in the PD brain. Exploratory immunoblotting analyses indicated hyperacetylation of multiple histone sites on histones H2B, H3 and H4, with the strongest change observed for H3K27. Importantly, this hyperacetylation was not due to changes in cellular composition of the samples and is not likely to be induced by anti-Parkinson drugs. The observed upregulation of the histone deacetylases SIRT1 and SIRT3 corroborates previous observations[41] and is compatible with decreased sirtuin activity caused by the presumed NAD^+ deficiency in PD[25,26,41].

Our immunoblot findings prompted us to map the genomic landscape of H3K27ac in the PD prefrontal cortex. ChIP-seq analysis corroborated our immunoblot results, showing an increase in the total number of reads in peaks, and pointed out to genome-wide H3K27 hyperacetylation in PD. This is supported by multiple observations: i) When the H3K27ac peaks were called separately for PD and control samples, the PD group exhibited increased total number of peaks, increased percentage of unique peaks and increased genome coverage. ii) In both cohorts, higher read count in PD samples was found in most of the genomic regions. iii) The differential H3K27 acetylation between PD and controls is strongly associated with the effective gene length. The association is positive for hyperacetylated

regions and negative for hypoacetylation regions. This is what would be expected in case of general predisposition for hyperacetylation, which by chance is more likely to affect genes occupying larger genomic space.

In addition to the genome-wide distribution of the hyperacetylated H3K27 mark, some specific sites appear to be altered in a more targeted manner. Out of 75 genes previously associated with PD that were represented in our ChIP-seq data, 22 exhibited adjusted metaP value < 0.05 and 5 of the genes: *DLG2*, *MAP4K4*, *CRHR1*, *SH3GL2* and *MBNL2* were identified as DARs in each cohort independently. Thus, our findings support the role of these genes as the causal genes within the identified loci. Another gene of interest is *PTPRH*, encoding protein tyrosine phosphatase receptor type H, which is robustly hypoacetylated in our data. While there is not much known about the specific functions of this gene, it was previously associated with PD based on whole exome sequencing analysis[35]. Interestingly, the authors demonstrated that RNAi targeting of the homolog of this gene in *Drosophila*, enhanced α-synuclein toxicity[35].

The observed decoupling between promoter H3K27ac and gene expression in PD samples is highly intriguing in light of the known strong association between promoter H3K27 acetylation and gene expression[39]. While at this stage, we can only speculate about the exact nature of this observation, we have excluded the possibility that this is an artifact caused by our normalization method since the decoupling was observed irrespective of whether we applied the house-keeping or RLE normalization approach. It is also unlikely that the observed decoupling is due to lower RNA quality of the PD samples, since similar results were observed in two independent cohorts of subjects.

In summary, we provide the first insight into the genomic landscape of histone acetylation in PD. Our data show that H3K27 acetylation is severely dysregulated and decoupled from gene expression in the prefrontal cortex of individuals with idiopathic PD. Future studies in

purified and single cells will be crucial to understand whether H3K27 hyperacetylation represents a global or a cell-type specific phenomenon. Further research is warranted in order to elucidate the molecular mechanisms underlying aberrant histone acetylation in PD. Moreover, since histone acetylation can be modified by drugs modulating the action of histone acetylases and deacetylases, our results open the possibility that agents targeting histone acetylation may be exploited therapeutically in PD.

Online Methods

Cohorts

Prefrontal cortex (Brodmann area 9) tissue was obtained from two independent cohorts. The discovery cohort (ParkWest, or PW cohort) comprised individuals with idiopathic PD (n = 17) from the ParkWest study, a prospective population-based cohort which has been described in detail [23], and neurologically healthy controls (Ctrl, n = 15) from the Neuromics Biobank for Aging and Neurodegeneration. For a subset of the PW cohort (PD: n =7, Ctrl: n = 7), tissue samples from the striatum (putamen) and cerebellar cortex were also included. All individuals of the PW cohort had undergone whole-exome sequencing [42] and individuals with pathogenic mutations in genes implicated in Mendelian PD and other monogenic neurological disorders had been excluded. All PW individuals were of Western Norwegian origin.

The replication cohort, comprised individuals with idiopathic PD (n = 10) and neurologically healthy controls (n = 11) of Dutch origin, from the Netherlands Brain Bank (NBB, https://www.brainbank.nl/). Demographic information and experimental allocation of all samples is shown in Supplementary table 1. For each experiment, there were no significant differences for sex, age, or PMI between the compared groups.

None of the study participants had disorders or used medication known to influence histone acetylation status, during the last year of life. A complete list of the medication used by the individuals with PD during at least the last 12 months before death can be found in Supplementary Table 1. Individuals with PD fulfilled the National Institute of Neurological Disorders and Stroke [43] and the U.K. PD Society Brain Bank [44] diagnostic criteria for PD at their final visit. Controls had no clinical or neuropathological evidence of neurodegeneration, or other neurological disease.

Brain tissue

Brains were collected at autopsy and split sagittally along the corpus callosum. One hemisphere was fixed whole in formaldehyde and the other coronally sectioned and snap-frozen in liquid nitrogen. All samples were collected using a standard technique and fixation time of ~2 weeks. There were no significant differences in PMI between patient and control groups in each experiment (Supplementary Table 1). Routine neuropathological examination including immunohistochemistry for α-synuclein, tau and beta-amyloid was performed on all brains. All cases showed neuropathological changes consistent with PD including degeneration of the dopaminergic neurons of the *substantia nigra pars compacta* in the presence of Lewy pathology. Controls had no pathological evidence of neurodegeneration.

Chemicals, reagents and media

Unless otherwise indicated, all chemicals were from Sigma Aldrich or from Bio-Rad. Cell culture media and reagents were from Invitrogen.

Cell culture

SH-SY5Y cells (ATCC® CRL-2266™) were cultivated in DMEM /F-12 (1:1) medium with GlutaMAX containing 10 % FBS and penicillin / streptomycin. For differentiation, 8 x 10^5 cells were seeded in a 6-well plate, and at about 80 % confluency sequentially treated with 10 µM retinoid acid in complete medium for 5 days and serum-free medium containing 50 ng / ml BDNF (Tocris) for 3 days as described previously [45]. For pharmacological treatment, fully differentiated cells were exposed to commonly prescribed anti-Parkinson drugs (APDs) including L-dopa (50 µM), ropinirole (0.05 µM), entacapone (0.05 µM and 0.5 µM) and carbidopa (30 µM and 1.5 µM,) for 48 h, or deacetylase inhibitor trichostatin A (TSA, 0.5 µM) for 8 hours. The applied concentrations were calculated based on the maximum clinical dose according to the Norwegian prescription catalogue

(https://www.felleskatalogen.no/medisin). The maximum prescribed dosage was converted to blood concentration, according to reported pharmacokinetic properties in the Norwegian prescription catalogue, and used in the experiment. When the calculated values were below the described IC50 of the compound, an additional sample using the IC50 concentration was included. For L-DOPA, the concentration was chosen based on previous reports [46]. Cell viability was assessed by resazurin assay to optimize drug exposure time. Each cell culture experiment was conducted three times starting with undifferentiated SH-SY5Y cells.

Lysate preparation, protein determination and immunoblotting

Fresh frozen brain tissue was lysed in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA and 2 % SDS. PBMCs and SH-SY5Y cells were lysed in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.2 % SDS, 10 µM trichostatin A, 10 µM nicotinamide and complete protease inhibitor cocktail. All samples were homogenized with a 27G syringe and centrifuged at 12.000 x g for 5 min at room temperature. Supernatants were collected and protein concentration determined using the BCA protein kit (Pierce/Thermofisher). Three independently dissected samples of brain tissue were analyzed from each individual. Immunoblots were repeated at least 3 times for each marker.

For immunoblotting, 15 µg total protein were separated using AnykD™ MiniProtean® TGX™ Precast Protein gradient gels (Bio-Rad) and transferred to a 0.2 µm PVDF membrane using the Trans-Blot Turbo system (Bio-Rad) at 25 V and 1 Amp for 30 min. After transfer, membranes were blocked with 5 % BSA in TBS containing 0.1 % Tween-20 (TBS-Tween) for antibodies targeting acetylated lysines and in 5 % nonfat drymilk in TBS-Tween for all other antibodies. Incubation with primary antibody was carried out overnight at 4 ^{o}C. Primary antibodies were used against acetylated lysine (CST #9441, 1:10.000), histone H2AK5ac (Abcam, ab1764, 1:2000), histone H2BK15ac (CST #9083, 1:10.000), histone H3K9K14ac

(CST #9677, 1:20.000), histone H3K27ac (CST #8173S, 1:1000), histone H3K56ac (CST #4243, 1:200), histone H4K5ac (CST #8647, 1:1000), histone H4K12ac (CST #13944S, 1:2000), histone H4K16ac (CST #13534, 1:5000), histone H2A (Abcam, ab18255, 1:1000), histone H2B (Abcam, ab1790, 1:10.000), histone H3 (Abcam, ab1791m 1:20.000), histone H4 (CST # 2935, 1:500), GAPDH (SantaCruz, sc-32233, 1:20.000), NeuN (Millipore, MAB377, 1:1000), GFAP (Sigma-Aldrich, G3893, 1:15.000), CNP1 (Atlas antibodies, AMAb91072, 1:4000), CX3CR1 (Abcam, ab8021, 1:500), NUP188 (Invitrogen, PA5-48940, 1:250), calbindin (Sigma-Aldrich C9848, 1:3000), α-tubulinK40ac (Abcam, ab179484, 1:2000), α-tubulin (Abcam, ab18251, 1:2000), SIRT1 (Sigma Aldrich, S5447, 1:250), SIRT2 (Sigma Aldrich, S8447, 1:2000), SIRT3 (Sigma Aldrich, S4072, 1:2000).

After washing four times with TBS-Tween for 5 min, HRP-conjugated secondary antibodies (rabbit anti-mouse (P0260) or swine anti-rabbit (P0217), DAKO) were incubated for one hour at RT. After washing three times with TBS-Tween and once with TBS for 5 min, immunodetection was carried out using enhanced chemiluminescence (Clarity ECL detection kit, Bio-Rad), images were taken on a ChemiDoc XRS imager (Bio-Rad) and bands were quantified by densitometry using the ImageLab software (Bio-Rad).

To normalize the histone acetylation signal to histone quantity, each acetylation blot was stripped twice using stripping buffer containing 0.2 M glycine, 3.5 mM SDS and 1 % Tween-20 at pH 2.2 for one hour at 50 ^{o}C. Complete stripping was confirmed by incubation of the membranes with secondary antibodies only and ECL. No bands were detected even after long exposure time (up to 10 min). After detection, the membrane was washed and re-probed with primary antibody recognizing the histone being studied for assessment of total histone protein amount. To exclude potential bias due to incomplete stripping, we performed additional experiments where acetylated and total histone markers for each histone modification were assessed in parallel on different blots. To estimate the acetylated fraction of each histone

marker, we calculated the signal ratio of acetylated marker to total histone protein. This ratio is referred to as the "acetylation fraction" for each histone marker.

Immunoblot detection for each acetylation marker was repeated at least three times, each time using a freshly prepared tissue lysate sample of the same individual. Data from repeated runs retrieved by densitometry measurement for both the acetylation ratio and the total histone were assessed by repeated measures ANOVA that showed no significant differences within runs. Group comparison was done by linear mixed models using disease status, age, sex and PMI as fixed effects.

Statistical analysis of WB data

Comparison of WB densitometric values between PD and controls was performed using linear regression for panacetylation and linear mixed-models for all other markers. Linear mixed models analysis was chosen in order to account for intraindividual random variability. Subjects were specified as random effects and disease state, sex, age of death, PMI and cell markers were used as fixed effects. The error distribution was assumed normal, and solution to the mixed models was obtained using the maximum likelihood estimation method. The mixed-effect model analyses were implemented in SPSS version 22.0.0.

Chromatin Immunoprecipitation sequencing (ChIP-Seq)

Frozen brain tissue was shipped to Active Motif where the ChIP-sequencing experiment was performed. For each ChIP-seq experiment approximately 200mg of frozen brain tissue per sample was fixed in PBS + 1% formaldehyde at room temperature for 15 minutes. Fixation was stopped by the addition of glycine to final concentration 0.125 M. The tissue pieces were then treated with a TissueTearer and finally spun down and washed 2x in PBS. Tissue was then homogenized with a lysis buffer using a manual Dounce homogenizer. Lysates were

sonicated and the DNA sheared to an average length of 300-500 bp. 14 ug of chromatin was precleared with protein A agarose beads (Invitrogen). Protein-DNA complexes were immuno-precipitated using 4 ug of H3K27Ac antibody (Active Motif, cat# 39133, lot# 01518010). Complexes were washed, eluted from the beads with SDS buffer, followed by de-crosslinking with proteinase K and RNase treatment overnight at 65 C. The ChIP DNA was purified by phenol-chloroform extraction and ethanol precipitation. Quantitative PCR (qPCR) reactions were carried out in triplicate on specific genomic regions using SYBR Green Supermix (Bio-Rad). The resulting signals were normalized for primer efficiency by carrying out qPCR for each primer pair using Input DNA. Illumina sequencing libraries were prepared from the ChIP and Input DNAs by the standard consecutive enzymatic steps of end-polishing, dA-addition, and adaptor ligation. After a final PCR amplification step, the resulting DNA libraries were quantified and sequenced on Illumina's NextSeq 500 (75 base pair reads, single end).

Sequencing files QC and filtering

Raw FASTQ files were trimmed using Trimmomatic version 0.38 [47] with options ILLUMINACLIP:TruSeq2-SE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:25. Assessment of FASTQ files was carried out with FastQC version 0.11.8 [http://www.bioinformatics.babraham.ac.uk/projects/fastqc] before and after trimming. Reads were aligned to the GRCh37 human genome using bwa version 0.7.17 [48]. Reads with mapping quality < 30 were discareded using samtools version 1.9 [http://www.htslib.org/] and duplicated reads removed using MarkDuplicates from Picard Tools version 2.18.27 [http://broadinstitute.github.io/picard/]. BAM files were converted into BED files and sorted.

Peak calling

Peak calling procedure was carried out in two complementary approaches: 1) **Per group (PD or controls):** the peaks are called separately for each group, aggregating the relevant samples and inputs. While this approach is not optimal for differential peak calling analysis as it requires harmonization of peaks between the two groups, it can be used to compare the total genome coverage as well as unique H3K27ac genomic regions in each group. 2) **All samples combined:** the peaks are called by aggregating all samples and input controls. This approach does not allow identification of unique genomic regions however, it provides a better identification of consensus peaks, and thus, was used for differential peak analysis.

Specifically, BED files for single samples were merged into the three groups: (a) all PD samples, (b) all control samples (c) all samples. MACS2 version 2.1.1 [49] was used on each of the three group BED files without the shifting model (--nomodel) and with the following parameters: sequencing tag size (-s) 75, read extension (--extsize) 200, read shift (--shift) 0, and keeping all duplicates (--keep-dup all). The three resulting peak-sets were filtered using the ENCODE black listed regions and only peaks within canonical chromosomes were kept for downstream analyses. Sample-specific enrichment on the identified peaks was performed using featureCounts version 1.6.4 [50] with default parameters. To this end, reads within filtered BAM files were extended to reach the fragment length (200) before quantification. For both the individual group and all samples combined approaches, peaks with non-adjusted p-value > 10^{-7} were excluded from the analysis.

Calculation of Marker Site Profiles (MSP)

In order to identify brain cell-type specific H3K27ac regions we first analyzed H3K27ac ChIP-seq data from $NeuN^+$ and $NeuN^-$ brain cells [28]. Celltype based broadPeaks (CellType_peak-set), BAM files and metadata files were downloaded from https://www.synapse.org/#!Synapse:syn4566010. Differential regions between $NeuN^+$

(neurons) and NeuN$^-$ (glia) cells were calculated using "DESeq2" R package, including chromatin amount, library batch, sex, hemisphere, age and pH as covariates in the model. Peaks were defined as cell-type specific differentially acetylated regions (DARs) if they met the following criteria: 1) |fold change| > 4 and 2) mean count > 1000. Peaks were annotated to genes using the "build_annotations" function from "annotatr" R package based on UCSC hg19 genome assembly. Peaks were annotated to all genes for which they intersected a region between 5kb upstream from their TSS to the end of their 5'UTR. In the next step, we intersected the genes with DARs between glia and neurons with expression-based cortical marker gene lists based on NeuroExpresso database [30]. The DARs were next reassigned to specific glial and neuronal cells if they were annotated to genes defined as cell-type specific based on NeuroExpresso. For example, all DARs annotated to *MBP*, defined as oligodendrocyte marker gene based on NeuroExpresso, were defined as oligodendrocyte marker sites (MSS). In the same way, all DARs annotated to *SLC1A3*, defined as astrocyte marker gene based on NeuroExpresso, were defined as astrocyte MSS. In the next step, we used HTSeq to quantify the reads inside CellType_peak-set in our samples. The values were then converted to counts per million (CPM) and transformed to log2(CPM + 1). Next, for each cell-type specific MSS, we performed principal component analysis based on the relevant peaks using "prcomp" function from the "stats" R package using (scale = T), as described in [30]. Marker Site Profiles (MSP) were defined as the scores of the samples in the first principal component, transformed to {0-1} range for visualization purposes.

Identification of differentially acetylated regions

In order to further increase the reliability of analyzed peaks, for differential peaks, we removed peaks with maximal coverage of < 5 reads/base, calculated as max($Peak_{counts}$)/$Peak_{length}$). Housekeeping normalization factor was calculated as follows: For

each peak annotated to a house-keeping gene with mean count > 150 (housekeeping peak), for each sample, we first calculated the ratio between the sample count to the mean count of the peak across all samples ($counts_{sample}/counts_{mean}$). The housekeeping normalization factor was then defined as mean of the ratios for all relevant housekeeping peaks. Housekeeping normalization factors were calculated based on two sets of housekeeping genes: 1) Housekeeping genes suggested based on stable transcript expression [27] and 2) Housekeeping genes manually selected based on known high and global expression level and very high and broad peak signal covering most of the gene body.

Differential peak analysis was performed using "nbinomWaldTest" function from "DESeq2" R package. Dispersions were calculated by using "estimateDispertions" function from "DESeq2" R package, by setting the fitType parameter to "local". Normalization factors calculated by the "estimateSizeFactors" function were replaced with the housekeeping normalization factor based on our manually selected housekeeping gene set (*ACTB, GAPDH, UBC*). Peaks were annotated to genes as described in the previous section. The output of the analysis was obtained using DESeq2 "results" function, setting the cutoff for adjusted p-value to 0.05, and independentFiltering = T.

Association of GWAS-hits and EGL with observed altered H3K27 acetylation

Effective gene length was defined as sum of length of all regions annotated to the gene (including 1Kb upstream to TSS). We assessed the association between altered H3K27 acetylation defined as -log10(metaP) with being annotated to a GWAS gene, EGL and the total number of common regions aligned to the gene using a linear model. The total number of common regions in a gene was included as a covariate in the model. Namely, for a common region i, annotated to gene j, the -log(metaP) value was modeled as: $-\log10(metaP)_{ij} \sim GWAS_j + \log10(EGL_j) + \sum CommonRegions_j$. For the analysis was performed in each cohort

separately, metaP was replaced with cohort-specific p-value, and the total number of common regions was replaced with the total number of regions (peaks) annotated to the gene in the cohort-specific peak-set.

Promoter H3K27ac - expression correlation

For each gene, ChIP-seq/RNA-seq seq counts were adjusted for the covariates used for the differential analysis and Pearson's correlation was computed using "cor.test" R function, in each group (controls or PD subjects) separately. Only peaks annotated to promoter regions of the genes were included in the analysis. In cases where multiple promoter peaks were defined for a gene, the correlation was calculated for each promoter separately, and the only the highest correlation was kept. This scenario was either a result of multiple promoters due to existence of several splice variants, or in rare cases, multiple peaks in the same promoter. The choice of the highest correlation was performed independently for each group. Since not all genes are expected to exhibit a good correspondence between promoter H3K27ac state and expression level due to existence of multiple levels of regulation, we compared the distribution of correlations in each group for different thresholds of absolute correlation in either of the groups (e.g., cor > |0.1| means that only genes with absolute correlation > 0.1 in at least one of the groups are included in the analysis). Group comparison was performed using Wilcoxon rank sum test, using "wilcox" R function.

Code availability

The R code and the data can be made available upon request to the corresponding author.

Acknowledgments

This work is supported by grants from The Research Council of Norway (ES633272) and Bergen Research Foundation (BFS2017REK05).

The cell-type specific H3K27ac ChIP-seq data were generated as part of the PsychENCODE Consortium, supported by: U01MH103392, U01MH103365, U01MH103346, U01MH103340, U01MH103339, R21MH109956, R21MH105881, R21MH105853, R21MH103877, R21MH102791, R01MH111721, R01MH110928, R01MH110927, R01MH110926, R01MH110921, R01MH110920, R01MH110905, R01MH109715, R01MH109677, R01MH105898, R01MH105898, R01MH094714, P50MH106934, U01MH116488, U01MH116487, U01MH116492, U01MH116489, U01MH116438, U01MH116441, U01MH116442, R01MH114911, R01MH114899, R01MH114901, R01MH117293, R01MH117291, R01MH117292 awarded to: Schahram Akbarian (Icahn School of Medicine at Mount Sinai), Gregory Crawford (Duke University), Stella Dracheva (Icahn School of Medicine at Mount Sinai), Peggy Farnham (University of Southern California), Mark Gerstein (Yale University), Daniel Geschwind (University of California, Los Angeles), Fernando Goes (Johns Hopkins University), Thomas M. Hyde (Lieber Institute for Brain Development), Andrew Jaffe (Lieber Institute for Brain Development), James A. Knowles (University of Southern California), Chunyu Liu (SUNY Upstate Medical University), Dalila Pinto (Icahn School of Medicine at Mount Sinai), Panos Roussos (Icahn School of Medicine at Mount Sinai), Stephan Sanders (University of California, San Francisco), Nenad Sestan (Yale University), Pamela Sklar (Icahn School of Medicine at Mount Sinai), Matthew State (University of California, San Francisco), Patrick Sullivan (University of North Carolina), Flora Vaccarino (Yale University), Daniel Weinberger (Lieber Institute for Brain Development), Sherman Weissman (Yale University), Kevin White (University of Chicago), Jeremy Willsey (University of California, San Francisco), and Peter Zandi (Johns Hopkins University.

We are grateful to Prof. Laurence A. Bindoff and Dr. Shreejoy Tripathy for the valuable comments and discussions.

Contributions

L.T. conceived, designed and performed the ChIP-Seq and integrative computational analyses. G.T.T. designed and performed the immunoblots and cell-culture experiments. J.S. contributed to conception and implementation of the ChIP-seq and immunoblot analyses. K.H. contributed to data interpretation and critical revision of the manuscript. G.S.N. performed the preprocessing of the ChIP-seq dataset and contributed to the computational analyses. C.D. conceived, designed, supervised the immunoblots and cell-culture experiments, performed part of the immunoblot experiments, contributed to data analyses and interpretation. O.B.T. and G.A. contributed part of the tissue material, supplementary data, and critical revision of the manuscript. C.T. conceived, designed and directed the study, contributed to data analyses and interpretation and acquired funding for the study. L.T.,

G.T.T., C.D. and C.T. wrote the manuscript with the active input and participation of J.S. and G.S.N. All authors have read and approved the manuscript.

Ethics declarations

Ethics considerations

Informed consent was available from all individuals. Ethical permission for these studies was obtained from our regional ethics committee (REK 2017/2082, 2010/1700, 131.04)

Competing Interests

The authors report no competing financial interests.

References

1. de Rijk, M. C. *et al.* Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology* **54**, S21-3 (2000).
2. Dickson, D. W. Parkinson's disease and parkinsonism: neuropathology. *Cold Spring Harb Perspect Med* **2**, (2012).
3. Schapira, A. H. *et al.* Mitochondrial complex I deficiency in Parkinson's disease. *Lancet* **1**, 1269 (1989).
4. Flones, I. H. *et al.* Neuronal complex I deficiency occurs throughout the Parkinson's disease brain, but is not associated with neurodegeneration or mitochondrial DNA damage. *Acta Neuropathol* (2017). doi:10.1007/s00401-017-1794-7
5. Klein, C. & Westenberger, A. Genetics of Parkinson's disease. *Cold Spring Harb Perspect Med* **2**, a008888 (2012).
6. Nalls, M. A. *et al.* Expanding Parkinson's disease genetics: novel risk loci, genomic context, causal insights and heritable risk. *bioRxiv* 388165 (2019). doi:10.1101/388165
7. Keller, M. F. *et al.* Using genome-wide complex trait analysis to quantify 'missing heritability' in Parkinson's disease. *Hum Mol Genet* **21**, 4996–5009 (2012).
8. Wirdefeldt, K., Gatz, M., Reynolds, C. A., Prescott, C. A. & Pedersen, N. L. Heritability of Parkinson disease in Swedish twins: a longitudinal study. *Neurobiol. Aging* **32**, 1923.e1–8 (2011).
9. Noyce, A. J. *et al.* Meta-analysis of early nonmotor features and risk factors for Parkinson disease. *Ann. Neurol.* **72**, 893–901 (2012).
10. Peleg, S., Feller, C., Ladurner, A. G. & Imhof, A. The Metabolic Impact on Histone Acetylation and Transcription in Ageing. *Trends Biochem. Sci.* **41**, 700–711 (2016).
11. Saha, R. & Pahan, K. HATs and HDACs in neurodegeneration: a tale of disconcerted acetylation homeostasis. *Cell Death Differ.* **13**, 539–550 (2006).

12. Nativio, R. *et al.* Dysregulation of the epigenetic landscape of normal aging in Alzheimer's disease. *Nat Neurosci* **21**, 497–505 (2018).
13. Marzi, S. J. *et al.* A histone acetylome-wide association study of Alzheimer's disease identifies disease-associated H3K27ac differences in the entorhinal cortex. *Nat. Neurosci.* **21**, 1618 (2018).
14. Harrison, I. F., Smith, A. D. & Dexter, D. T. Pathological histone acetylation in Parkinson's disease: Neuroprotection and inhibition of microglial activation through SIRT 2 inhibition. *Neurosci. Lett.* **666**, 48–57 (2018).
15. Park, G. *et al.* Regulation of Histone Acetylation by Autophagy in Parkinson Disease. *J Biol Chem* **291**, 3531–40 (2016).
16. Gebremedhin, K. G. & Rademacher, D. J. Histone H3 acetylation in the postmortem Parkinson's disease primary motor cortex. *Neurosci. Lett.* **627**, 121–125 (2016).
17. Verdin, E. NAD^+ in aging, metabolism, and neurodegeneration. *Science* **350**, 1208–1213 (2015).
18. Karamanlidis, G. *et al.* Mitochondrial Complex I Deficiency Increases Protein Acetylation and Accelerates Heart Failure. *Cell Metab.* **18**, 239–250 (2013).
19. Anderson, K. A., Madsen, A. S., Olsen, C. A. & Hirschey, M. D. Metabolic control by sirtuins and other enzymes that sense NAD+, NADH, or their ratio. *Biochim. Biophys. Acta* **1858**, 991–998 (2017).
20. Sauve, A. A. Sirtuin Chemical Mechanisms. *Biochim. Biophys. Acta* **1804**, 1591–1603 (2010).
21. Imai, S. & Guarente, L. It takes two to tango: NAD+ and sirtuins in aging/longevity control. *NPJ Aging Mech. Dis.* **2**, 16017 (2016).
22. Singh, P., Hanson, P. S. & Morris, C. M. Sirtuin-2 Protects Neural Cells from Oxidative Stress and Is Elevated in Neurodegeneration. *Park. Dis.* **2017**, (2017).

23. Alves, G. *et al.* Incidence of Parkinson's disease in Norway: the Norwegian ParkWest study. *J Neurol Neurosurg Psychiatry* **80**, 851–7 (2009).

24. Balan, I. S., Fiskum, G. & Kristian, T. Visualization and quantification of NAD(H) in brain sections by a novel histo-enzymatic nitrotetrazolium blue staining technique. *Brain Res.* **1316C**, 112 (2010).

25. Braidy, N. *et al.* Age related changes in NAD+ metabolism oxidative stress and Sirt1 activity in wistar rats. *PloS One* **6**, e19194 (2011).

26. Koltai, E. *et al.* Exercise alters SIRT1, SIRT6, NAD and NAMPT levels in skeletal muscle of aged rats. *Mech. Ageing Dev.* **131**, 21–28 (2010).

27. Allhoff, M., Seré, K., F. Pires, J., Zenke, M. & G. Costa, I. Differential peak calling of ChIP-seq signals with replicates with THOR. *Nucleic Acids Res.* **44**, e153–e153 (2016).

28. Girdhar, K. *et al.* Cell-specific histone modification maps in the human frontal lobe link schizophrenia risk to the neuronal epigenome. *Nat. Neurosci.* **21**, 1126–1136 (2018).

29. Oldham, M. C. *et al.* Functional organization of the transcriptome in human brain. *Nat. Neurosci.* **11**, 1271–1282 (2008).

30. Mancarci, B. O. *et al.* Cross-Laboratory Analysis of Brain Cell Type Transcriptomes with Applications to Interpretation of Bulk Tissue Data. *eNeuro* ENEURO.0212-17.2017 (2017). doi:10.1523/ENEURO.0212-17.2017

31. Titus, A. J., Gallimore, R. M., Salas, L. A. & Christensen, B. C. Cell-type deconvolution from DNA methylation: a review of recent applications. *Hum. Mol. Genet.* **26**, R216–R224 (2017).

32. Sun, W. *et al.* Histone Acetylome-wide Association Study of Autism Spectrum Disorder. *Cell* **167**, 1385-1397.e11 (2016).

33. Andrade-Moraes, C. H. *et al.* Cell number changes in Alzheimer's disease relate to dementia, not to plaques and tangles. *Brain J. Neurol.* **136**, 3738–3752 (2013).

34. von Bartheld, C. S., Bahney, J. & Herculano-Houzel, S. The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting. *J. Comp. Neurol.* **524**, 3865–3895 (2016).
35. Jansen, I. E. *et al.* Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing. *Genome Biol.* **18**, 22 (2017).
36. Wu, H.-C. *et al.* DLG2, but not TMEM229B, GPNMB, and ITGA8 polymorphism, is associated with Parkinson's disease in a Taiwanese population. *Neurobiol. Aging* **64**, 158.e1-158.e6 (2018).
37. Foo, J. N. *et al.* Genome-wide association study of Parkinson's disease in East Asians. *Hum. Mol. Genet.* **26**, 226–232 (2017).
38. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci.* **107**, 21931–21936 (2010).
39. Karlić, R., Chung, H.-R., Lasserre, J., Vlahoviček, K. & Vingron, M. Histone modification levels are predictive for gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2926–2931 (2010).
40. Nido, G. S. *et al.* Common gene expression signatures in Parkinson's disease are driven by changes in cell composition. *bioRxiv* 778910 (2019). doi:10.1101/778910
41. Pallàs, M. *et al.* Modulation of SIRT1 expression in different neurodegenerative models and human pathologies. *Neuroscience* **154**, 1388–1397 (2008).
42. Gaare, J. J. *et al.* Rare genetic variation in mitochondrial pathways influences the risk for Parkinson's disease: Mitochondrial Pathways In PD. *Mov. Disord.* **33**, 1591–1600 (2018).
43. Gelb, D. J., Oliver, E. & Gilman, S. Diagnostic criteria for Parkinson disease. *Arch. Neurol.* **56**, 33–39 (1999).
44. Ward, C. D. & Gibb, W. R. Research diagnostic criteria for Parkinson's disease. *Adv Neurol* **53**, 245–9 (1990).

45. Encinas, M. *et al.* Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells. *J. Neurochem.* **75**, 991–1003 (2000).
46. Ingman, K. *et al.* The effect of different dosing regimens of levodopa/carbidopa/entacapone on plasma levodopa concentrations. *Eur. J. Clin. Pharmacol.* **68**, 281–289 (2012).
47. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
48. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–95 (2010).
49. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
50. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma. Oxf. Engl.* **30**, 923–930 (2014).

Tables & Figures

Table 1. Top hypoacetylated and hyperacetylated regions identified as DARs in both PW and NBB cohorts.

Symbol	Region Annotation	CHR	Region Location	Peak Location (PW, NBB)	MetaP	LFC PW, NBB	PeakName PW, NBB	Proportion Significant
Common hypoacetylated regions								
PTPRH**	introns	19	55718568-55720750	55720288-55722414, 55717912-55719103	1.7E-07	-0.81, -0.94	Peak_20188, Peak_15767	3^{p}/4(0.8)
JUP	introns	17	39681526-39923630	39805764-39808133, 39807092-39808006	1.1E-05	-0.43, -0.99	Peak_53983, Peak_81536	5^{p}/10^{p} (0.5)
Common hyperacetylated regions (top 15)								
DLG2*	exons	11	83874504-83874554	83873551-83879844, 83873666-83877544	7.7E-12	0.68, 1.05	Peak_21808, Peak_40741	30^{p}/49^{p}(0.6)
KIT	exons	4	55604595-55606881	55606747-55609274, 55606782-55609553	8.1E-11	0.63, 1.26	Peak_21777, Peak_37103	2/3^{p} (0.7)
CNTNAP2	exons	7	147336198-147336398	147333973-147338260, 147333944-147337213	6.2E-09	0.51, 1.07	Peak_44999, Peak_47588	11/18^{p} (0.6)
RNF216	introns	7	5681008-5692043	5679685-5685127, 5688370-5705729	1.2E-08	0.6, 0.53	Peak_38631, Peak_12618	11^{p}/16 (0.7)
FRMD5	introns	15	44216508-44393656	44260324-44263821, 44244249-44247825	1.6E-08	0.68, 0.67	Peak_18018, Peak_18744	12^{p}/17^{p}(0.7)
SOX2-OT	introns	3	181328718-181417385	181358880-181360150, 181341154-181348014	1.7E-08	0.55, 0.91	Peak_43667, Peak_31549	9/12(0.8)
MAGI2	1to5kb	7	78120338-78124337	78121257-78128843, 78115190-78121170	2.2E-08	0.4, 0.93	Peak_18816, Peak_24042	9^{p}/23(0.4)
PPP2R2B	introns	5	146080706-146460621	146383635-146384203, 146459075-146469815	2.8E-08	1.3, 0.52	Peak_135972, Peak_17645	17^{p}/27^{p}(0.6)
KCNH1	introns	1	210948887-210970849	210966438-210969215, 210946491-210950181	3.8E-08	0.55, 1.0	Peak_51838, Peak_38212	9/18^{p} (0.5)
CDH1	introns	16	68835797-68842326	68840159-68841130, 68782065-68837136	4.8E-08	0.84, 0.37	Peak_66179, Peak_3389	5/7^{p} (0.7)
MAP4K4*	introns	2	102315001-102407181	102330963-102334475, 102367013-102375949	5.2E-08	0.54, 0.62	Peak_30093, Peak_9009	10^{p}/24^{p}(0.4)
METTL24	introns	6	110644076-110729536	110656829-110661962, 110685537-110688887	5.4E-08	0.48, 0.97	Peak_22739, Peak_27591	8/13^{p}(0.6)
PIP4K2A	introns	10	22898647-23003111	22934570-22952399, 22957818-22962517	5.6E-08	0.58, 0.52	Peak_22480, Peak_27954	9^{p}/18^{p}(0.5)
SIK3	introns	11	116827781-116968858	116825819-116835556, 116846923-116884577	5.8E-08	0.6, 0.49	Peak_21504, Peak_11226	13^{p}/16^{p}(0.8)
DOCK10	introns	2	225796386-225906968	225798681-225800264, 225845772-225849871	8.1E-08	0.61, 0.77	Peak_94440, Peak_22866	6/10^{p}(0.6)
Common hyperacetylated regions (additional GWAS hits)								
CRHR1*	introns	17	43707525-43750192	43726544-43728643, 43741886-43747152	2.0E-07	0.43, 0.83	Peak_74390, Peak_30944	6/14^{p}(0.4)
MBNL2*	introns	13	97873814-97927885	97919075-97922481, 97906667-97916758	1.2E-05	0.44, 0.44	Peak_29656, Peak_16279	14^{p}/22^{p}(0.6)

For hyperacetylated regions, the top 15 genes with DARs common to the two cohorts and common DARs in genes previously associated with PD are shown. The genes were ranked based on Fisher's meta p-value. The full list of genes is provided in Supplementary Table S6. ** - Previous association with PD based on WES; * - Previous association with PD based on GWAS; Region Location – genomic location (on the indicated chromosome) of the annotated region; Peak location – genomic location of the relevant peak from each cohort (upper – PW, lower - NBB); LFC – log of change compared to controls; Peak Name – the peak identifier in each cohort; Proportion significance – DARs/total common regions annotated to the gene; ^{p}- peaks annotated to promoter region of the gene. In bold, genes retaining common DARs with RLE normalization

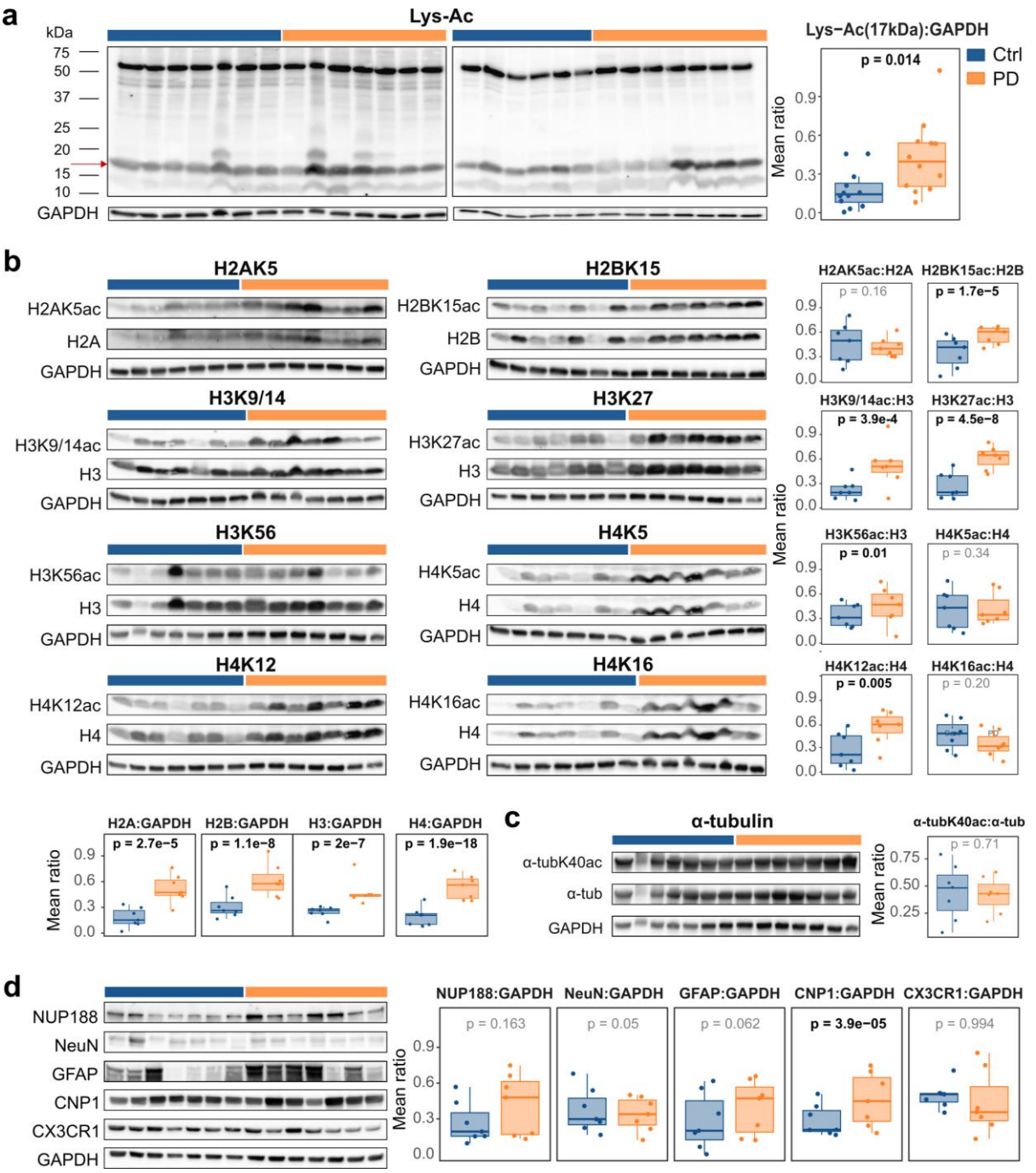


Fig. 1: Hyperacetylation of multiple histone sites in the PD prefrontal cortex

Representative Western blots and the ratios of their quantified densitometric measurements of the indicated proteins (or their modifications) in individuals with PD (orange) and controls (blue) are shown. GAPDH served as loading control. Values were rescaled 0-1. Boxplots show the mean of replicates for each individual. The indicated p-values are based on a linear mixed model (see methods). **a,** Global protein acetylation detected by a pan-acetyl lysine antibody. **b,** Relative acetylation of histone lysine residues and quantification of the total histone proteins. **c,** Relative acetylation of α-tubulin on lysine K40. **d,** Immunodetection and densitometric quantification of relative protein levels of the nuclear marker (NUP188) and cell type-specific marker proteins: NeuN – neurons, GFAP – astrocytes, CNP1 – oligodendrocytes, CX3CR1 - microglia.

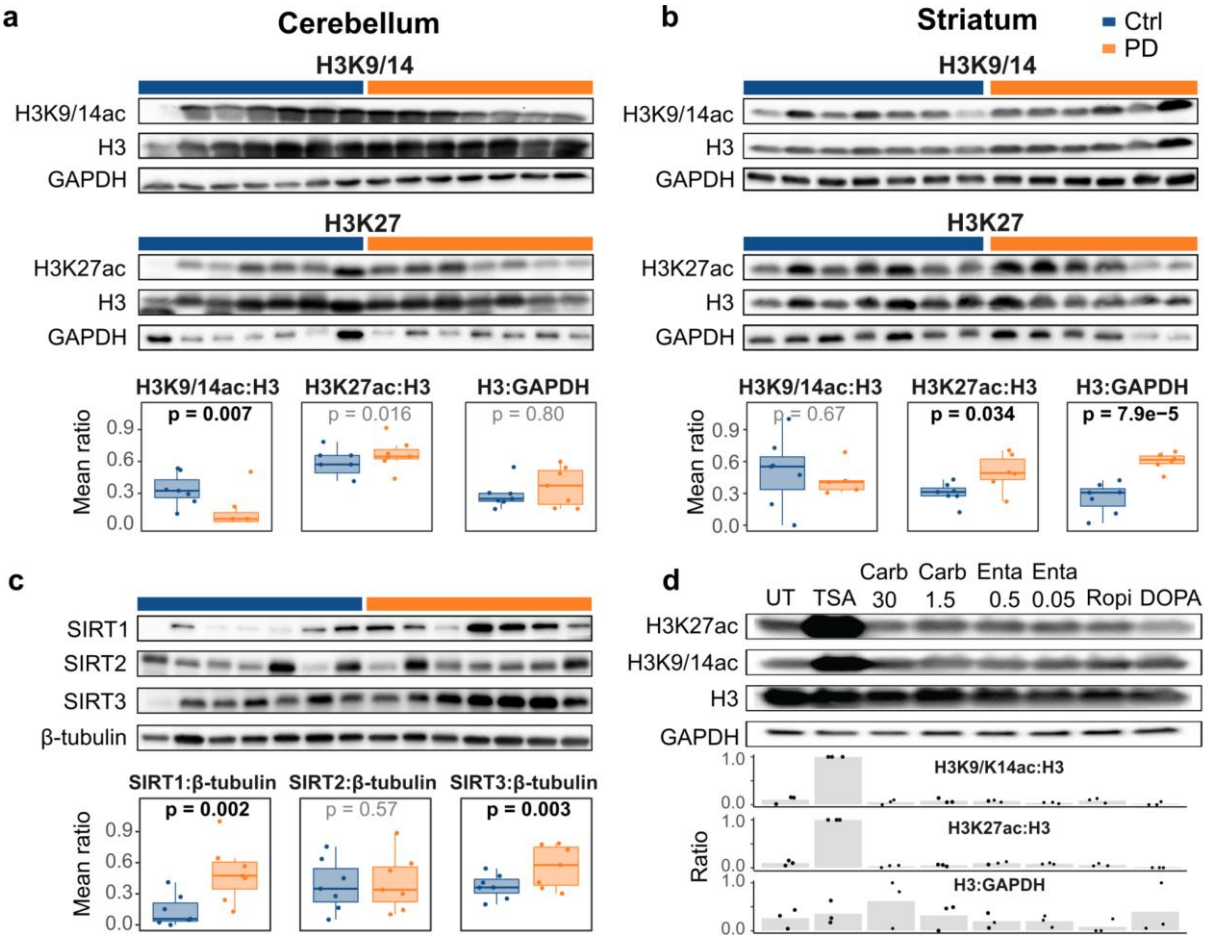


Fig. 2: Altered H3 acetylation occurs in multiple brain regions, is accompanied by sirtuin upregulation and is not induced by anti-Parkinson drugs in-vitro

a-c, Representative Western blots and quantification plots showing the ratios of quantified densitometric measurements of the indicated proteins in individuals with PD (orange) and controls (blue). GAPDH and β-tubulin serve as loading controls. Values were rescaled 0-1. Boxplots show the mean of replicates for each individual. The indicated p-values are based on a linear mixed model (see methods). **a, b,** Acetylated H3K9/14 and H3K27 and total histone protein H3 levels in the cerebellum (a) and striatum (b) from individuals with PD and controls. **c,** Sirtuin proteins in the prefrontal cortex of PD and controls. **d,** Acetylated H3K9/14, H3K27 and total H3 protein in differentiated SH-SY5Y-derived dopaminergic neuron-like cells treated with the histone deacetylase inhibitor Trichostatin A (TSA) or anti-Parkinson drugs used by the PW patients during their last year of life. The numbers indicate the concentration (in µM) of the drug. Carb: Carbidopa; Enta: Entacapone; Ropi: Ropinirole; DOPA: L-Dopa. UT: untreated control. A representative blot is shown. Bar plots show the mean quantified ratios of three biological replicates (rescaled 0-1). Dots indicate the values of each of the individual replicates.

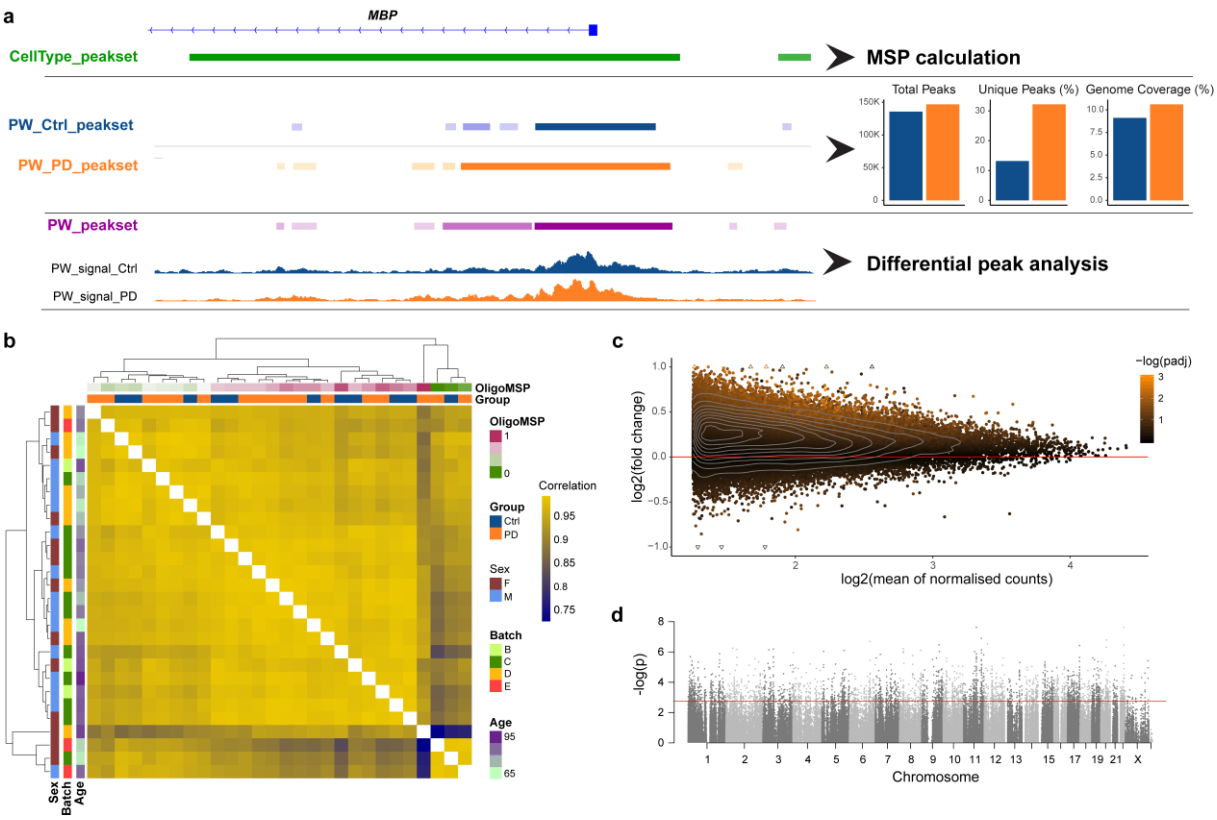


Fig. 3: H3K27ac ChIP-seq analysis of PD and control samples.
a, Different peak-sets used for the different stages of ChIP-seq analysis. Shown, as an example, is a genomic region around the oligodendrocyte marker *MBP*. Peaks annotated to this region, in each peak-set are shown by colored boxes. The CellType_peak-set was used for the MSP calculation. Group specific peak-sets were used to characterize the distribution of H3K27ac bound regions in cases and controls. Barplots show the comparison between the two groups (Supplementary Table S3). PD group exhibited higher number of peaks, higher percentage of unique peaks and higher percentage of genomic coverage, consistent with genome-wide H3K27 hyperacetylation. The PW_peak-set was used for the differential peak analysis. Shown are group overlays of sample fold of change compared to input control. **b,** Hierarchical clustering of the samples indicated that samples cluster based on their cellular composition. Supplementary Figure S6 shows the association of each of the variables with the main principal components of the data. **c**, MA plot based on the PW_peak-set indicates that increased H3K27ac in PD is observed genome-wide, rather than being restricted to specific regions. **d,** Manhattan plot showing the distribution of genomic locations and differential p-values of the H3K27ac PW_peak-set. The red dashed line indicates the -log10 of the highest p-value with < 5% false positive rate (Benjamini-Hochberg adjusted p-value < 0.05). Corresponding plots for the NBB cohort are shown in Supplementary Figure S5.

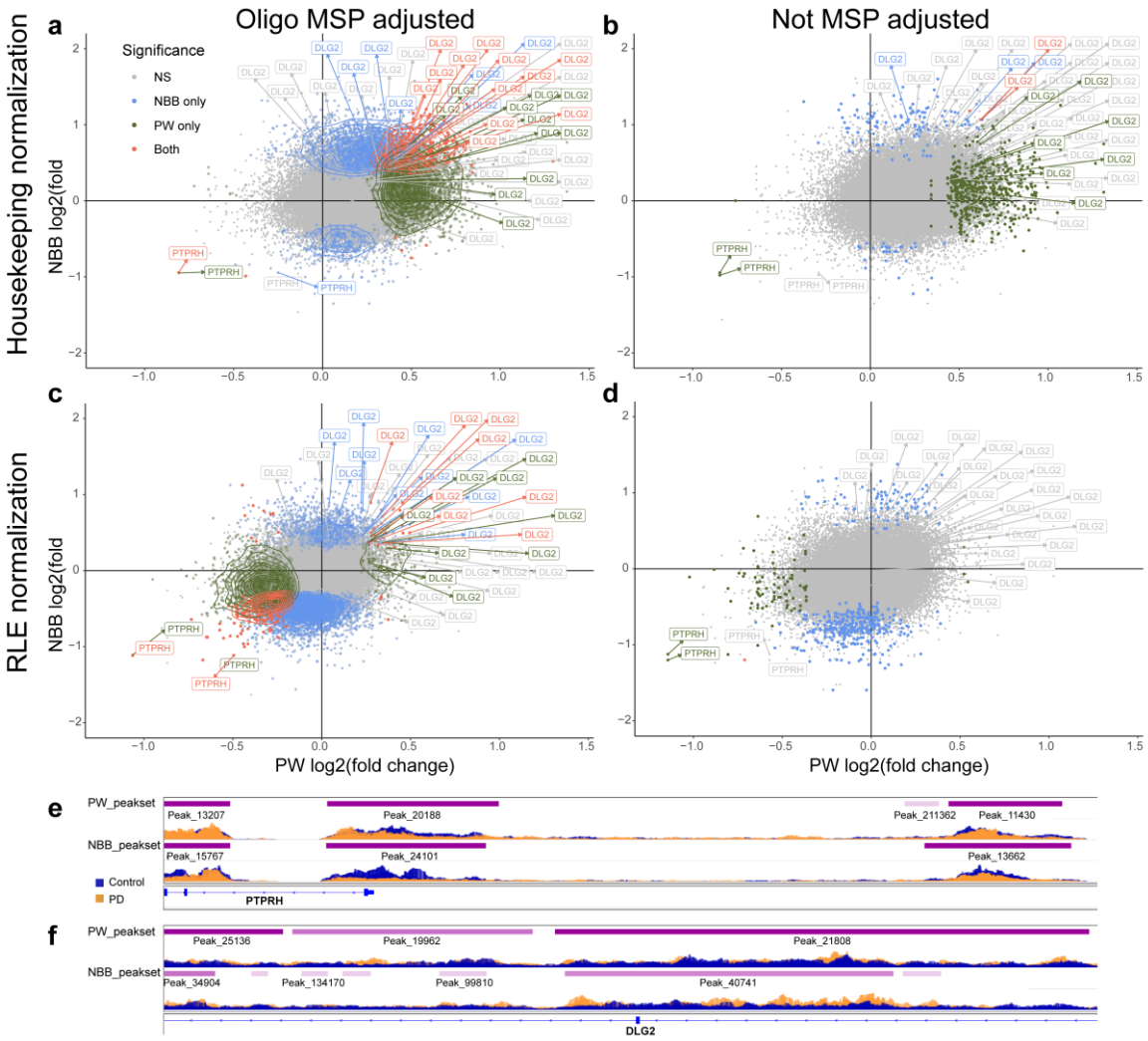


Fig. 4: Effect of normalization method and cell composition correction on H3K27ac ChIP-seq outcome

a-d, Log2 fold change of comparable H3K27ac regions from NBB (y axis) and PW (x axis) were plotted against each other. The color indicates whether and in which of the cohorts the region was identified as DAR. Both cell-type correction (**a,c** vs**. b,d**) and normalization method (**a,b** vs. **c,d**) have a dramatic effect on the outcome of the analysis. **e, f** – IGV tracks for two of the common DARs are shown for *PTPRH* (hypoacetylated, **e**) and *DLG2* (hyperacetylated, **f**) genes. Y-axis represents the overlay of fold of change relative to input control.

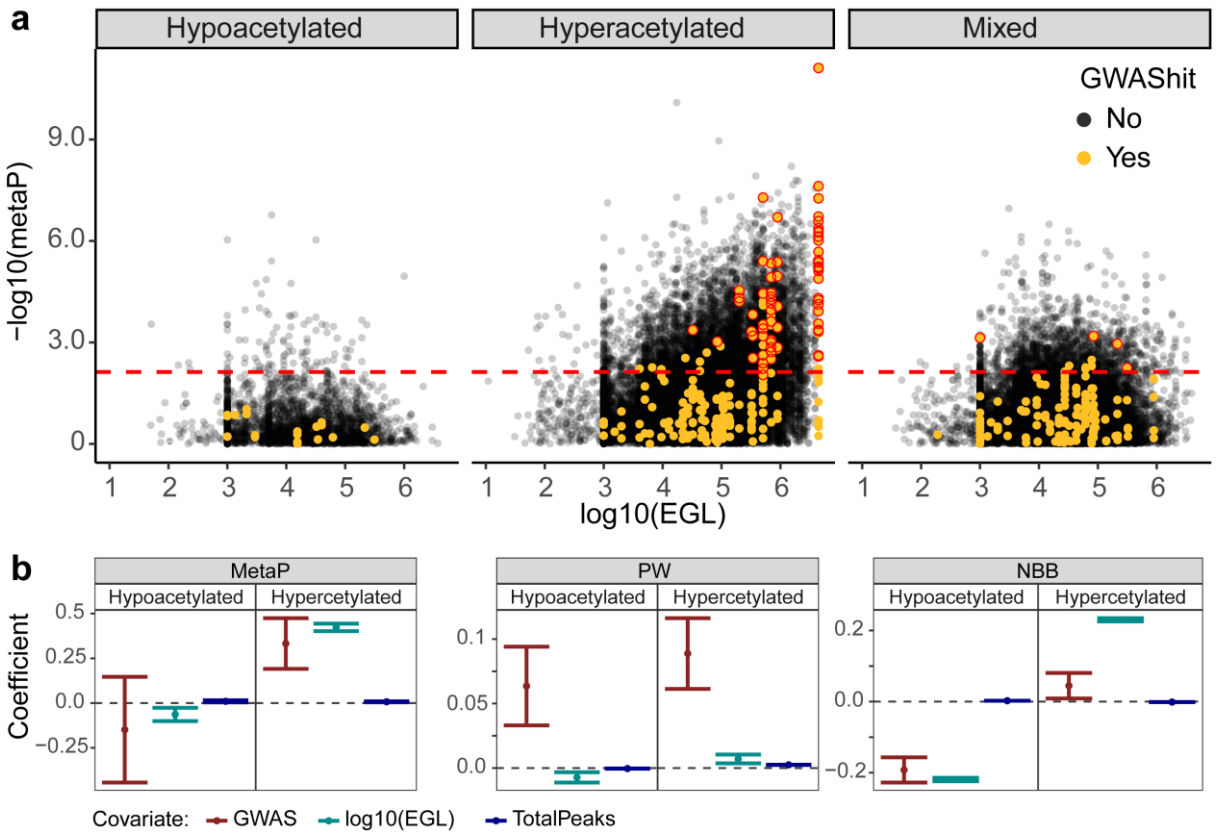


Fig. 5: The H3K27 acetylation state is strongly associated with PD-GWAS hits and the effective gene length

a, Common regions were divided into three groups: Hypoacetylated – regions showing hypoacetylation in both cohorts; Hyperacetylated – regions showing hyperacetylation in both cohorts; Mixed – regions with opposite acetylation trends in the two cohorts. The dashed line represents the -log10 of the highest metaP corresponding to adjusted metaP < 0.05. Multiple common regions within the same gene are shown. Highlighted are regions identified as DAR in at least one of the cohorts within PD-GWAS hits. metaP – Fisher's meta-p value of the corresponding two peaks from PW and NBB cohorts. **b,** Association between -log(p) and each of the indicated covariates based on linear model. Shown are the coefficients and 95% confidence intervals. The results are shown for the common regions (metaP) and each of the cohorts separately. Peaks annotated to miRNAs and snRNAs were excluded from the analysis. GWAS – a gene is a PD-GWAS hit. EGL – effective gene length. TotalPeaks – total number of common regions or peaks annotated to the gene.

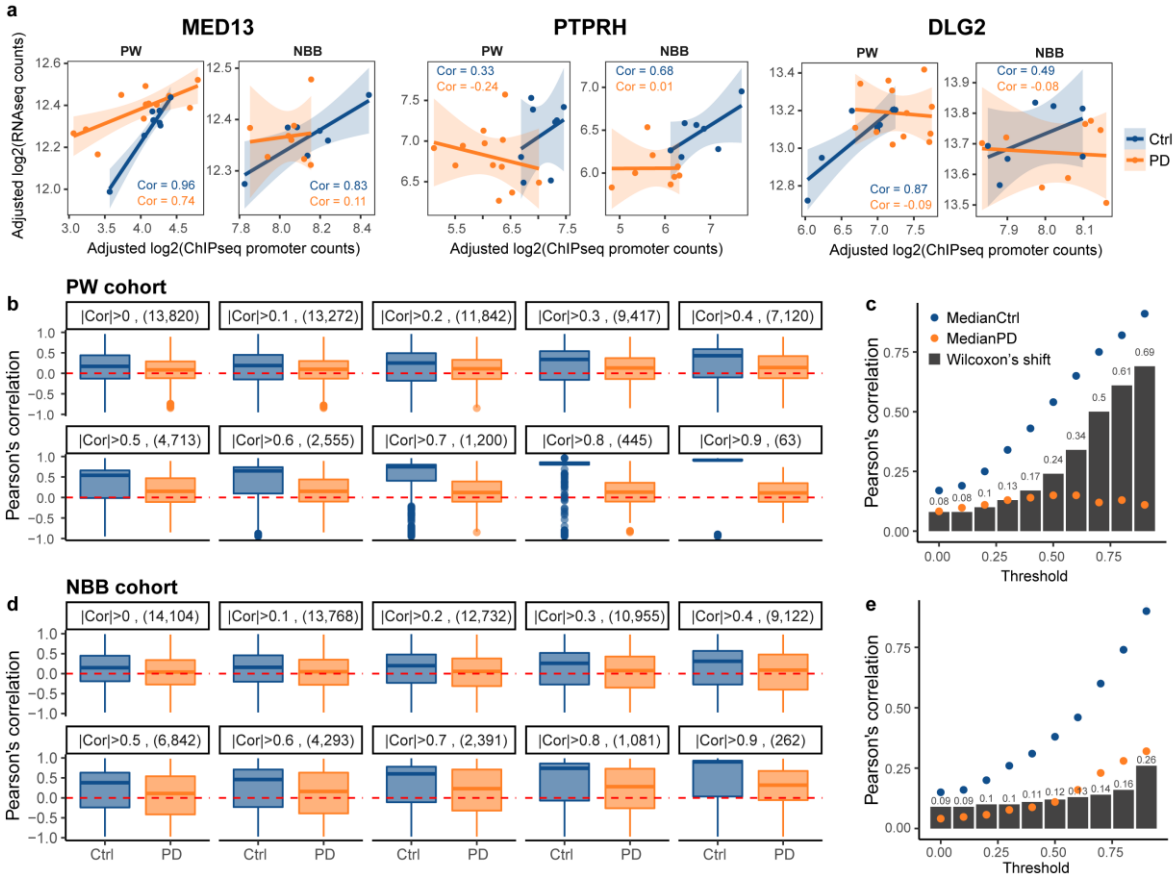
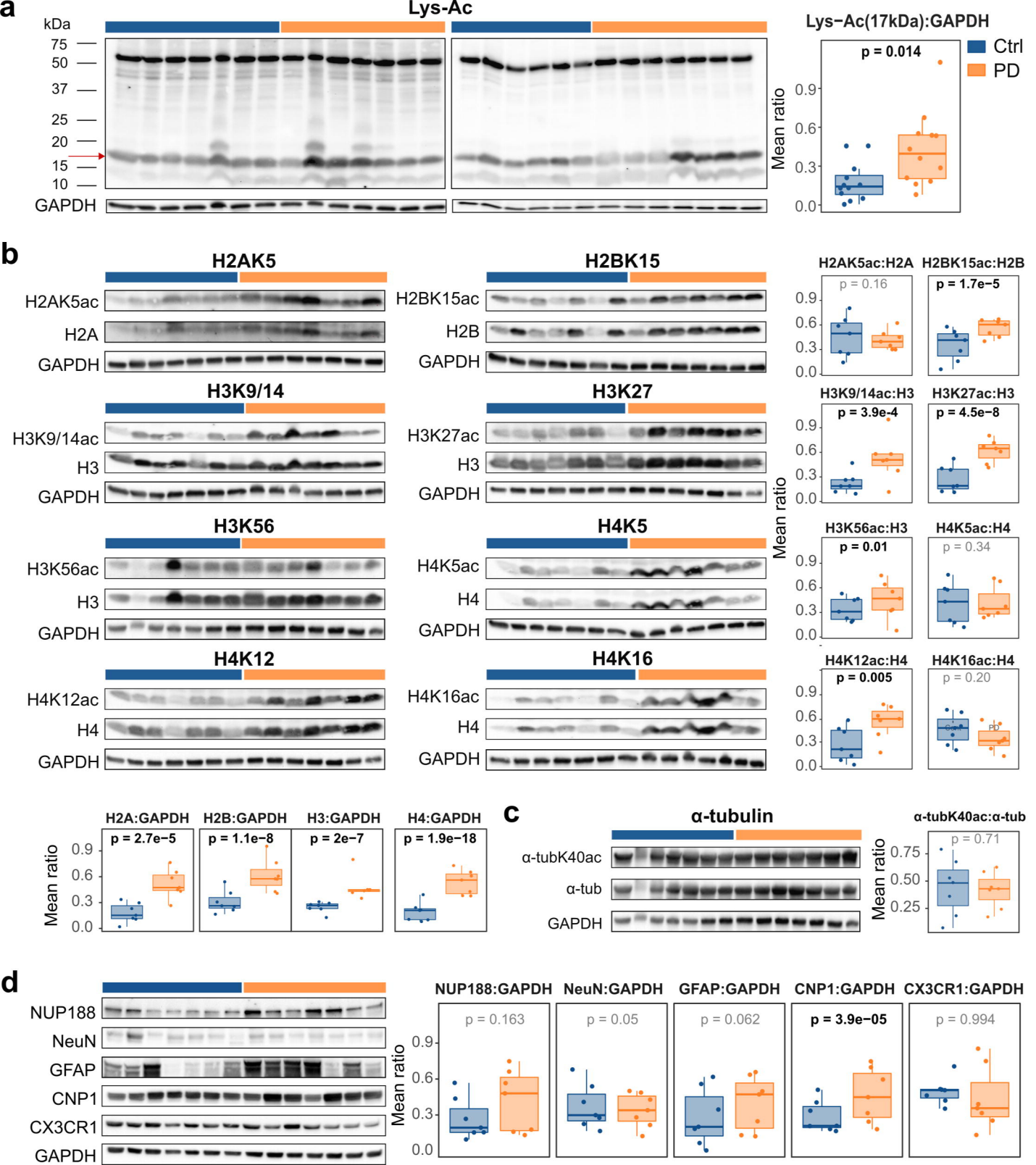
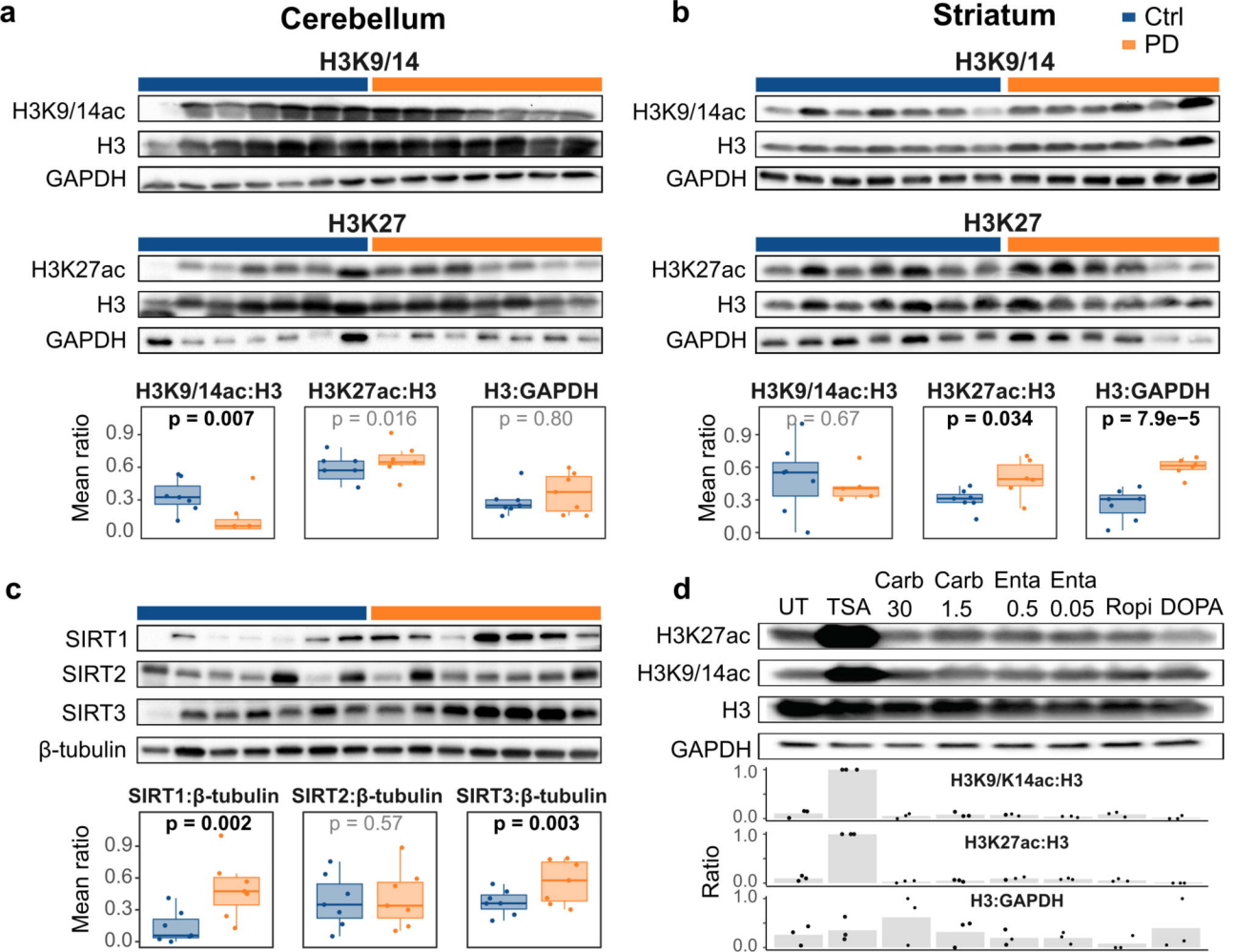
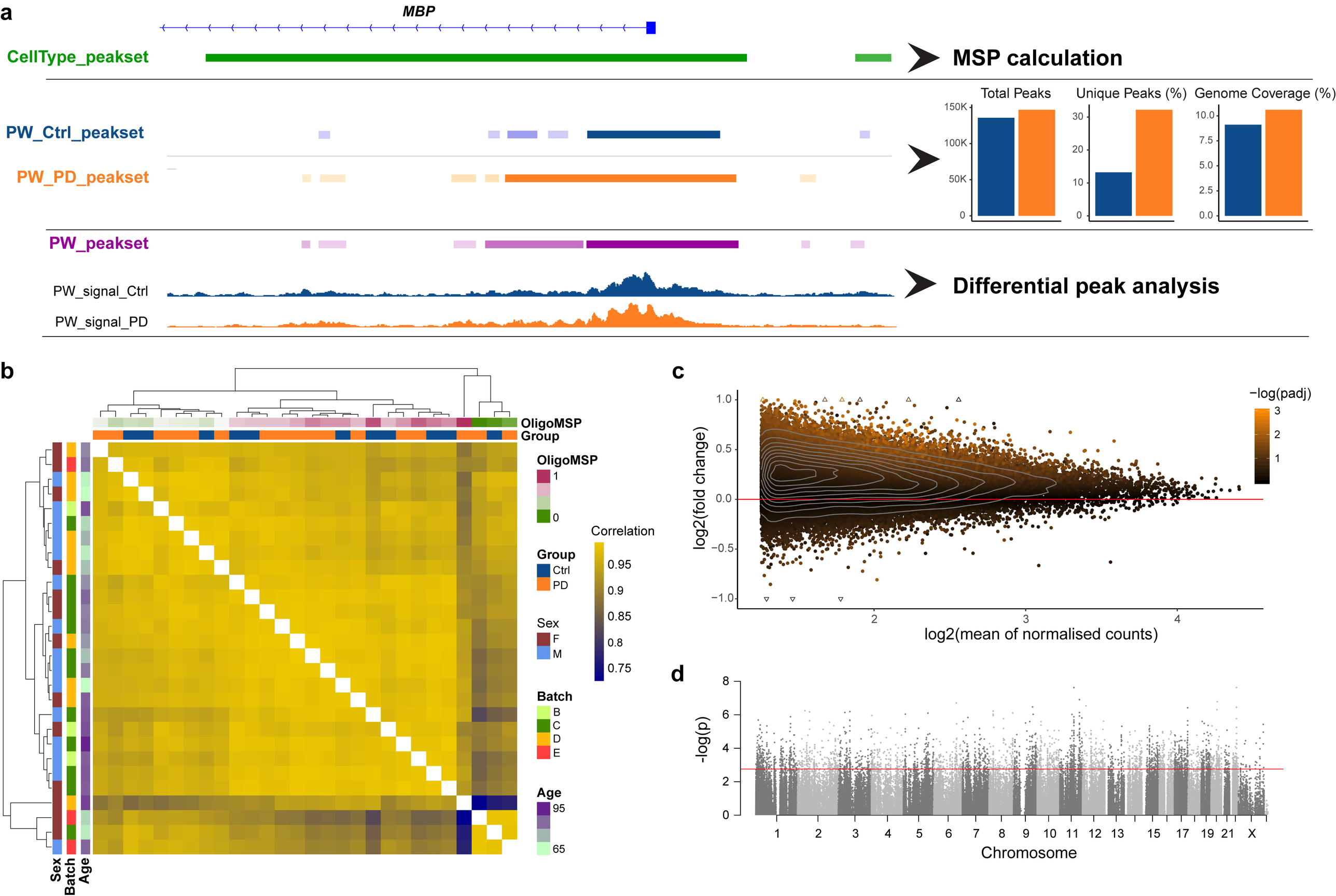


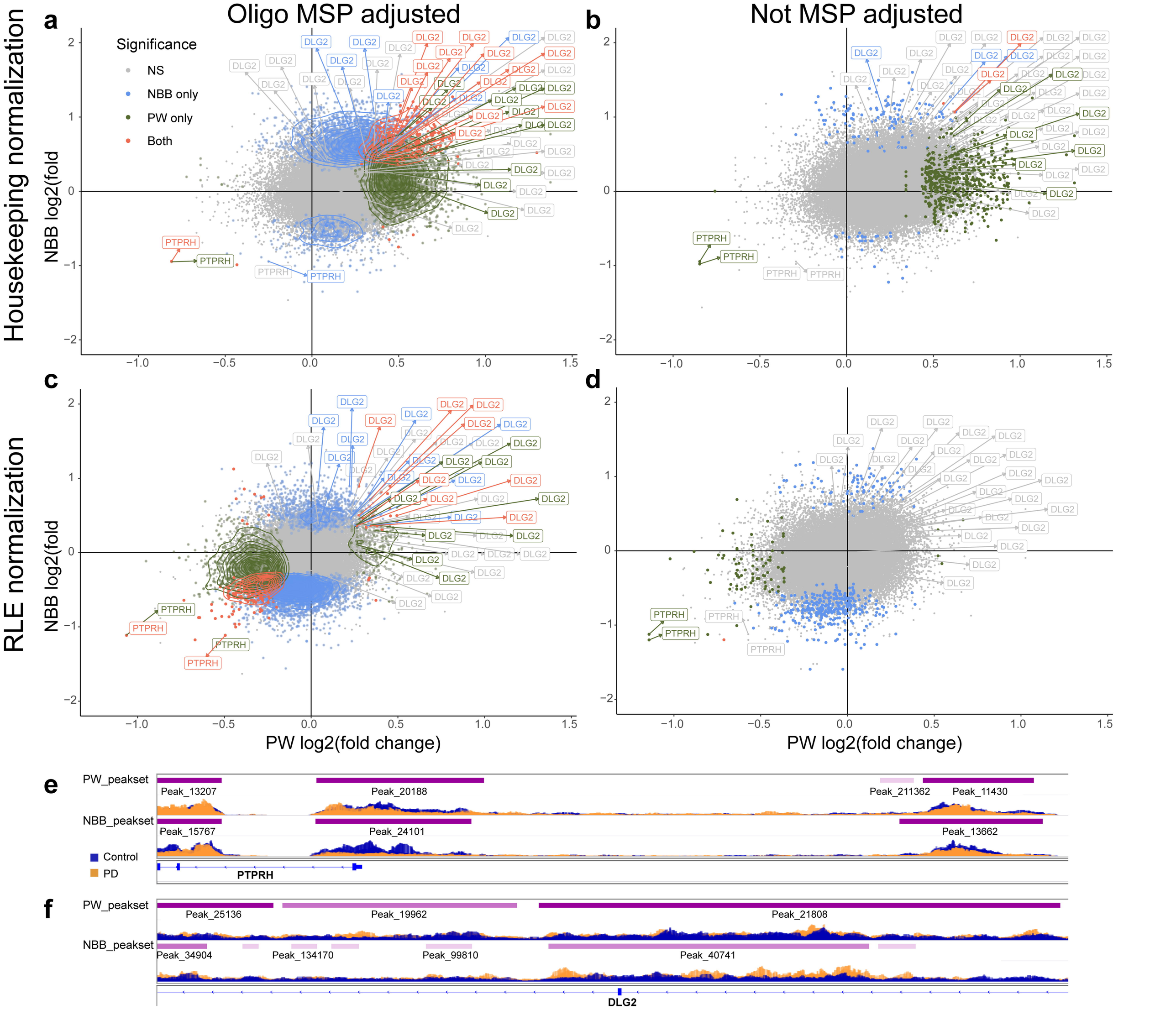
Fig. 6: Decreased correlation between promoter H3K27 acetylation state and gene expression in PD

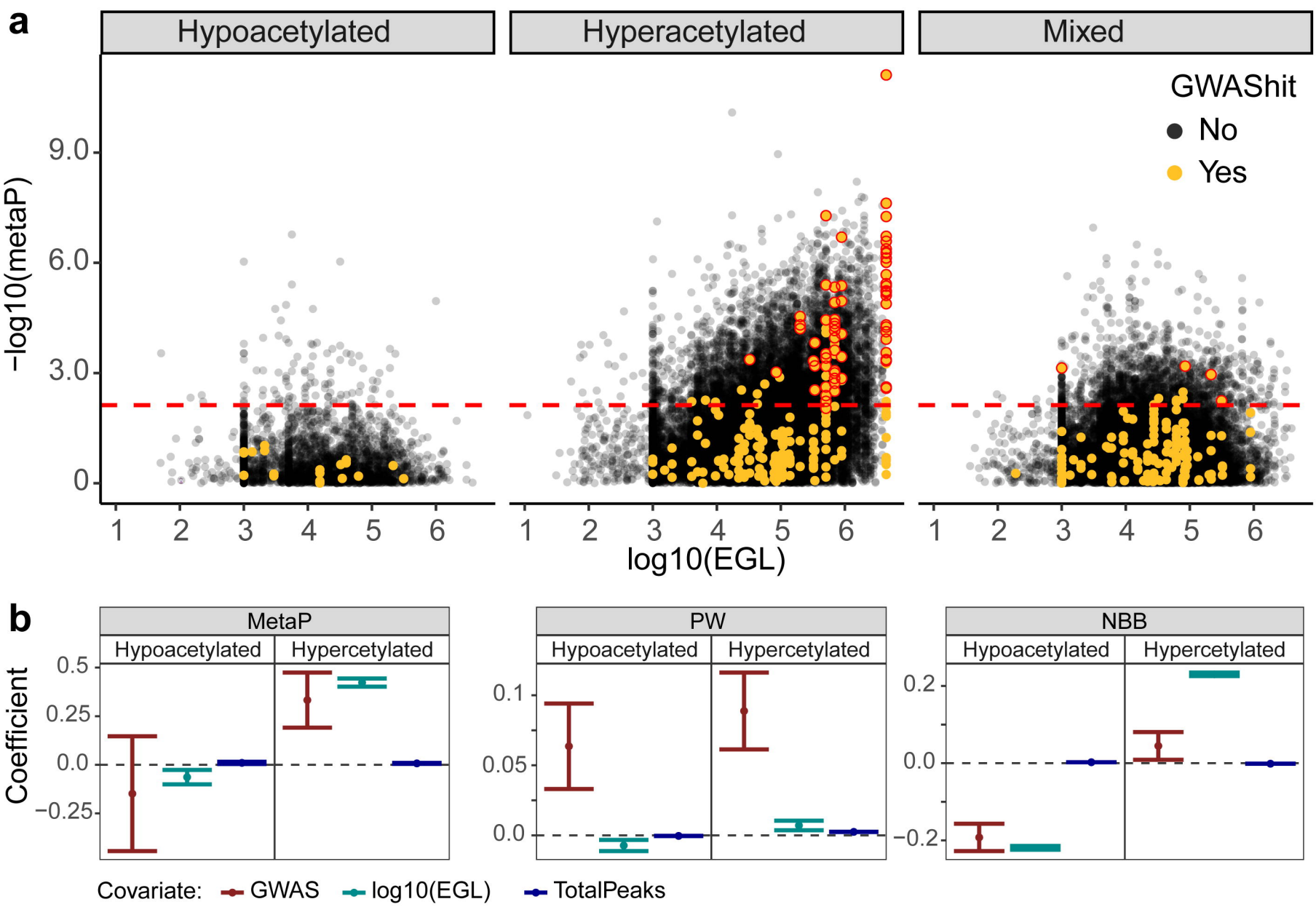
Pearson's correlation between the adjusted promoter H3K27ac counts and the adjusted RNA-seq counts was calculated for each gene across control or PD subjects in each cohort separately. **a,** Representative scatter plots showing the adjusted promoter H3K27ac counts (X-axis), the adjusted RNA-seq counts (Y-axis) and the calculated Pearson's correlation for three genes: *MED13*, *PTPRH* and *DLG2* in control (blue) or PD (orange) individuals from PW or NBB cohorts. Each dot represents an individual. Cor - Pearson's correlation. **b-e,** The distribution of correlations for all compatible genes was compared between the groups for various thresholds of minimal absolute correlation in either of the groups. Results are shown for PW (**b,c**) and NBB (**d,e**) cohorts. The number of genes that remained after applying the correlation threshold is shown in parentheses. The correlations remained closely distributed around 0 in PD subjects from both cohorts, regardless of the correlation threshold. **c,e,** Median correlations in each group and the calculated Wilcoxon's delta shift for the different correlation thresholds.

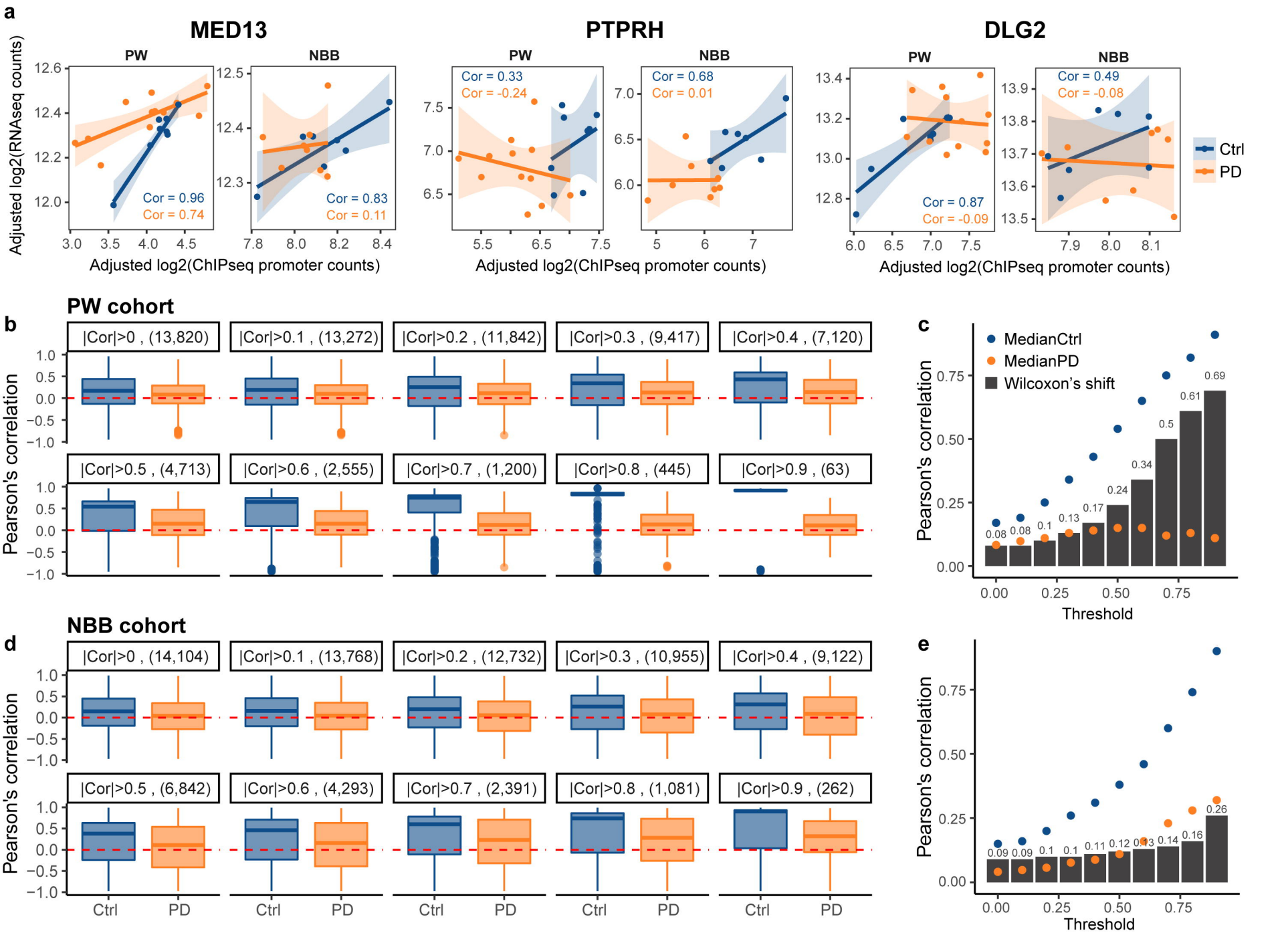

a
Lys-Ac
kDa
75
50
37
25
20
15
10
GAPDH
Lys−Ac(17kDa):GAPDH
p = 0.014
Ctrl
PD
Mean ratio
b
H2AK5
H2AK5ac
H2A
GAPDH
H2BK15
H2BK15ac
H2B
GAPDH
H3K9/14
H3K9/14ac
H3
GAPDH
H3K27
H3K27ac
H3
GAPDH
H3K56
H3K56ac
H3
GAPDH
H4K5
H4K5ac
H4
GAPDH
H4K12
H4K12ac
H4
GAPDH
H4K16
H4K16ac
H4
GAPDH
H2AK5ac:H2A
p = 0.16
H2BK15ac:H2B
p = 1.7e−5
H3K9/14ac:H3
p = 3.9e−4
H3K27ac:H3
p = 4.5e−8
H3K56ac:H3
p = 0.01
H4K5ac:H4
p = 0.34
H4K12ac:H4
p = 0.005
H4K16ac:H4
p = 0.20
H2A:GAPDH
p = 2.7e−5
H2B:GAPDH
p = 1.1e−8
H3:GAPDH
p = 2e−7
H4:GAPDH
p = 1.9e−18
c
α-tubulin
α-tubK40ac
α-tub
GAPDH
α-tubK40ac:α-tub
p = 0.71
d
NUP188
NeuN
GFAP
CNP1
CX3CR1
GAPDH
NUP188:GAPDH
p = 0.163
NeuN:GAPDH
p = 0.05
GFAP:GAPDH
p = 0.062
CNP1:GAPDH
p = 3.9e−05
CX3CR1:GAPDH
p = 0.994


a
Cerebellum
H3K9/14
H3K9/14ac
H3
GAPDH
H3K27
H3K27ac
H3
GAPDH
H3K9/14ac:H3
p = 0.007
H3K27ac:H3
p = 0.016
H3:GAPDH
p = 0.80
Mean ratio
0.9
0.6
0.3
0.0
b
Striatum
Ctrl
PD
H3K9/14
H3K9/14ac
H3
GAPDH
H3K27
H3K27ac
H3
GAPDH
H3K9/14ac:H3
p = 0.67
H3K27ac:H3
p = 0.034
H3:GAPDH
p = 7.9e−5
Mean ratio
0.9
0.6
0.3
0.0
c
SIRT1
SIRT2
SIRT3
β-tubulin
SIRT1:β-tubulin
p = 0.002
SIRT2:β-tubulin
p = 0.57
SIRT3:β-tubulin
p = 0.003
Mean ratio
0.9
0.6
0.3
0.0
d
UT
TSA
Carb 30
Carb 1.5
Enta 0.5
Enta 0.05
Ropi
DOPA
H3K27ac
H3K9/14ac
H3
GAPDH
H3K9/K14ac:H3
H3K27ac:H3
H3:GAPDH
Ratio
1.0
0.0


a
MBP
CellType_peakset
MSP calculation
PW_Ctrl_peakset
PW_PD_peakset
Total Peaks
Unique Peaks (%)
Genome Coverage (%)
150K
100K
50K
0
30
20
10
0
10.0
7.5
5.0
2.5
0.0
PW_peakset
PW_signal_Ctrl
PW_signal_PD
Differential peak analysis
b
OligoMSP
Group
OligoMSP
1
0
Group
Ctrl
PD
Sex
F
M
Batch
B
C
D
E
Age
95
65
Correlation
0.95
0.9
0.85
0.8
0.75
Sex
Batch
Age
c
−log(padj)
3
2
1
1.0
0.5
0.0
−0.5
−1.0
log2(fold change)
2
3
4
log2(mean of normalised counts)
d
8
6
4
2
0
-log(p)
1
2
3
4
5
6
7
8
9
10
11
12
13
15
17
19
21
X
Chromosome


Housekeeping normalization
RLE normalization
a
Oligo MSP adjusted
b
Not MSP adjusted
c
d
Significance
NS
NBB only
PW only
Both
NBB log2(fold
PW log2(fold change)
DLG2
PTPRH
e
PW_peakset
NBB_peakset
Peak_13207
Peak_20188
Peak_211362
Peak_11430
Peak_15767
Peak_24101
Peak_13662
Control
PD
PTPRH
f
PW_peakset
NBB_peakset
Peak_25136
Peak_19962
Peak_21808
Peak_34904
Peak_134170
Peak_99810
Peak_40741
DLG2


a
Hypoacetylated
Hyperacetylated
Mixed
GWAShit
No
Yes
−log10(metaP)
9.0
6.0
3.0
0
1
2
3
4
5
6
log10(EGL)
b
MetaP
PW
NBB
Hypoacetylated
Hypercetylated
Coefficient
0.5
0.25
0.0
−0.25
0.1
0.05
0.2
−0.2
Covariate:
GWAS
log10(EGL)
TotalPeaks


a
MED13
PTPRH
DLG2
PW
NBB
Cor = 0.96
Cor = 0.74
Cor = 0.83
Cor = 0.11
Cor = 0.33
Cor = -0.24
Cor = 0.68
Cor = 0.01
Cor = 0.87
Cor = -0.09
Cor = 0.49
Cor = -0.08
Ctrl
PD
Adjusted log2(RNAseq counts)
Adjusted log2(ChIPseq promoter counts)
PW cohort
b
|Cor|>0 , (13,820)
|Cor|>0.1 , (13,272)
|Cor|>0.2 , (11,842)
|Cor|>0.3 , (9,417)
|Cor|>0.4 , (7,120)
|Cor|>0.5 , (4,713)
|Cor|>0.6 , (2,555)
|Cor|>0.7 , (1,200)
|Cor|>0.8 , (445)
|Cor|>0.9 , (63)
Pearson's correlation
c
MedianCtrl
MedianPD
Wilcoxon's shift
Threshold
NBB cohort
d
|Cor|>0 , (14,104)
|Cor|>0.1 , (13,768)
|Cor|>0.2 , (12,732)
|Cor|>0.3 , (10,955)
|Cor|>0.4 , (9,122)
|Cor|>0.5 , (6,842)
|Cor|>0.6 , (4,293)
|Cor|>0.7 , (2,391)
|Cor|>0.8 , (1,081)
|Cor|>0.9 , (262)
e